

۱-

$$1. P_{N=5} = \binom{5}{3} \times 0.7^3 \times 0.3^2 + \binom{5}{4} \times 0.7^4 \times 0.3 + 0.7^5 = 0.83692$$

۵ بار درست
۴ بار درست
۳ بار درست

$$2. P_{N=9} = \binom{9}{5} \times 0.7^5 \times 0.3^4 + \binom{9}{6} \times 0.7^6 \times 0.3^3 + \binom{9}{7} \times 0.7^7 \times 0.3^2 + \binom{9}{8} \times 0.7^8 \times 0.3 + 0.7^9 \approx 0.9$$

3. با توجه به قضیه مرکز بزرگ n تغییر تعداد $i.i.d$ داریم :

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{n} \quad n \rightarrow \infty \quad S_n \sim N\left(\mu, \frac{\sigma^2}{n}\right) \rightarrow \text{توزیع نرمال}$$

$$\mu = E[X_i], \quad \sigma^2 = \text{Var}[X_i]$$

$$P_N = \sum_{i=\frac{N+1}{2}}^N \binom{N}{i} \times 0.7^i \times 0.3^{N-i} \rightarrow \text{توزیع درجه ۱ بزرگ} \rightarrow \mu = nP(X_i=1) = 0.7n$$

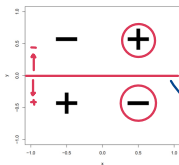
$$\sigma^2 = nP(X_i=1)P(X_i=0) = 0.21n$$

$$\xrightarrow{\text{قضیه حد مرکزی}} \lim_{n \rightarrow \infty} \Pr\{\sqrt{n}(S_n - \mu) \leq z\} = \Phi\left(\frac{z}{\sigma}\right) \rightarrow P_N \rightarrow 1$$

$\hookrightarrow N(0,1)$

$$4. P_{N=5} = \binom{5}{3} \times 0.5^3 \times 0.5^2 + \binom{5}{4} \times 0.5^4 \times 0.5 + 0.5^5 = 0.5 \rightarrow \text{شنای جمع random است}$$

در نتیجه اگر چند classifit با دقت ۵۰٪ داشته باشیم و بزرگ classification از جمع آنها استفاده کنیم به دقت بیشتر خواهیم رسید. اما اگر دقت آنها بیش از ۵۰٪ باشد (مثلاً ۶۰٪) می‌تواند با جمع آنها به دقت بالاتر از دقت classifit برسیم.



۱. الف)

Decision Stump یک دقت بهای با حسن است در طبقه بندی feature های که از استفاده کند
استفاده از هر feature یک دقت 1/50 می کند
بنابراین در classifier نهایی
حالا 1/50 مثلاً را تغییر ندهد
$$\epsilon_1 = 0,2 \rightarrow \alpha_1 = \frac{1}{2} \ln \frac{1 - \epsilon_1}{\epsilon_1} = 0$$

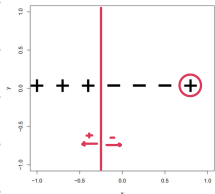
$$w_i^{(2)} = w_i^{(1)} \frac{e^{-\alpha_1 y_i h_1(x_i)}}{z_1} = w_i^{(1)} \rightarrow$$

2.
$$F_T(x) = \sum_{t=1}^T \alpha_t h_t(x) = F_{T-1}(x) + \alpha_T h_T(x)$$

$$\begin{aligned} E &= \sum_{i=1}^m e^{-y_i (F_{T-1}(x_i) + \alpha_T h_T(x_i))} = \sum_{i=1}^m \underbrace{e^{-y_i F_{T-1}(x_i)}}_{w_i^{(T-1)}} \cdot e^{-y_i \alpha_T h_T(x_i)} \\ &= e^{\alpha_T} \sum_{i \in M} w_i^{T-1} + e^{-\alpha_T} \sum_{i \in C} w_i^{T-1} = e^{\alpha_T} \epsilon_T + e^{-\alpha_T} (1 - \epsilon_T) \end{aligned}$$

$$\rightarrow \alpha_T = \frac{1}{2} \ln \left(\frac{1 - \epsilon_T}{\epsilon_T} \right) \rightarrow E = \epsilon_T \sqrt{\frac{1 - \epsilon_T}{\epsilon_T}} + (1 - \epsilon_T) \sqrt{\frac{\epsilon_T}{1 - \epsilon_T}} = 2 \sqrt{\epsilon_T (1 - \epsilon_T)} \quad ; \quad 0 < \epsilon < 0,5$$

3. ب)



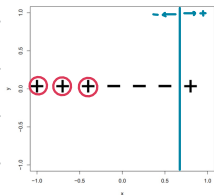
Decision Stump classifier ابتدا دقت هر داده را با یکسان در نظر می گیرد:

4.
$$\epsilon_1 = \frac{\sum_{i \in M} w_i^{(T-1)}}{\sum_{i=1}^m w_i^{(T-1)}} = \frac{1}{7} \rightarrow \alpha_1 = 0,396$$

5.
$$w_i^{(2)} = w_i^{(1)} \frac{e^{-\alpha_1 y_i h_1(x_i)}}{z_1} = \frac{0,0583}{z_1} = 0,0833 \quad ; \quad i \in \{1, \dots, 6\} \quad , \quad w_7^{(2)} = \frac{0,35}{z_1} = 0,5$$

داده ۷ را استفاده طبقه بندی شده

حزب منطقی هتم (که توسط h_1 به اشتباه classifed شد)، $\frac{1}{2}$ است. احتمال طبقه‌بندی آن نیز
 misclassified است. بالاست. در نتیجه از تقسیم به ترتیب در برهه خنجره برده:



7.
$$e_2 = \frac{\sum_{i \in A} w_i^{(2-1)}}{\sum_i w_i^{(2-1)}} = 0.0833 \times 3 = 0.25 \rightarrow e_2 = 0.55$$

$$w_i^{(3)} = w_i^{(2)} \frac{e^{-a_2 y_i h_2(x_i)}}{z_2}$$

$i = 1, 2, 3 \rightarrow 0.833 \times \frac{e^{0.55}}{z_2} = 0.1667$
 $i = 4, 5, 6 \rightarrow 0.833 \times \frac{e^{-0.55}}{z_2} = 0.05546$ (نویز کمتر - توسط h_2 درست طبقه‌بندی شده)
 $i = 7 \rightarrow 0.5 \times \frac{e^{-(0.144)}}{z_2} = 0.3333$

8. در مرحله دوم دست طبقه‌بندی کاهش می‌یابد. اما باعث می‌شود وزن داده‌ای که به اشتباه طبقه‌بندی شده‌اند، بالاتر برود. در مرحله غریب ربط به مشارکت h_2 هم
 (e_2) کاهش می‌یابد. که باعث می‌شود خطا در مجموع خیلی زیاده نشود.

5.

شماره	فشار خون X_1	سطح کلسترول X_2	مصرف سیگار X_3	وزن X_4	استاد شریان Y
۱	بله	۰	نه	۰	افاده وزن ۱
۲	نه	۰	بله	۰	نرمال
۳	نه	۲	نه	۰	افاده وزن ۱
۴	۰	۱	بله	۱	افاده وزن ۱
۵	بله	۲	بله	۱	جای ۲
۶	بله	۱	بله	۰	نرمال
۷	نه	۱	نه	۲	جای ۲
۸	بله	۰	بله	۰	نرمال
۹	بله	۲	نه	۰	جای ۲
۱۰	نه	۰	نه	۰	افاده وزن ۱
۱۱	نه	۲	بله	۱	نرمال
۱۲	بله	۱	نه	۱	افاده وزن ۱
۱۳	بله	۰	نرمال	۱	افاده وزن ۱
۱۴	بله	۱	بله	۲	جای ۲

استاد اطلاعات مورد نیاز را به ما. Information gain = مقدار منتقل می‌کند:

$$I(Y, X_i) = H(Y) - H(Y|X_i) \quad (\text{الف})$$

$$H(Y) = - \sum_Y P(Y=y) \log P(Y=y)$$

$$H(Y|X_i) = - \sum_n P(X_i=n) \sum_Y \frac{P(Y=y|X_i=n)}{\log P(Y=y|X_i=n)}$$

$$H(Y) = 0.9402859586706311$$

$$I(Y, X_1) = 0.04812703040826949$$

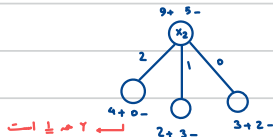
$$I(Y, X_2) = 0.24674981977443933$$

$$I(Y, X_3) = 0.15183550136234159$$

$$I(Y, X_4) = 0.02922256565895487$$

استاد از کدام مقدار

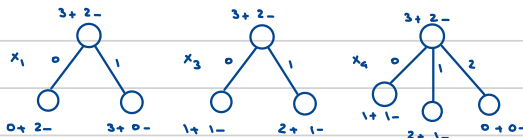
Information gain را بدست می‌آید:



ب. توجه: مقدار Information gain برای X_2 برابر ۰.۲۴۶۷۴۹۸۱۹۷۷۴۴۳۹۳۳ می‌باشد.

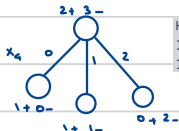
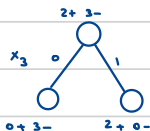
به این معنی که مقدار Information gain برای X_2 برابر ۰.۲۴۶۷۴۹۸۱۹۷۷۴۴۳۹۳۳ می‌باشد.

توجه: $X_2=0$ داریم:



بدین معنی که Info gain برای X_1 برابر ۰.۰۴۸۱۲۷۰۳۰۴۰۸۲۶۹۴۹ می‌باشد.

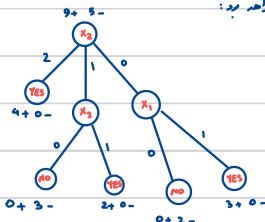
تک $x_2 = 1$ داریم (ب. ترجمه انتخاب x_1 در $x_2 = 0$ ، این ویژگی را برمی‌داریم):



$H(X_2) = 0.8772776505803251$
 $I(X_2, X_3) = 0.8772776505803251$
 $I(X_2, X_4) = 0.4772776505803251$

ب. ترجمه به عبارتی دیگر: استفاده از چهار x_3 بفرست.

در نهایت درخت تصمیم به شکل زیر خواهد بود:



شماره	فشار خون x_1	سطح کلسترول x_2	مصرف سیگار x_3	وزن x_4	انستداد شرایین	نتیجه پیش بینی
۱۵	بله	۱	نرمال	۰	بله	۱
۱۶	بله	۱	بالا	۱	چاق	۱
۱۷	بله	۱	بالا	۱	نرمال	۰
۱۸	بله	۱	نرمال	۰	نرمال	۱
۱۹	نه	۰	نرمال	۱	افزاده وزن	۰

حالا ماتریس اشتباهی را رسم می‌کنیم:

ب. ترجمه به ماتریس اشتباهی داریم:

	کلاس پیش بینی شده	
	۲	۱
کلاس واقعی	۲	۱
	۱	۱

$$Recall = \frac{TP}{TP + FN} = \frac{2}{3} = 0.66$$

$$Precision = \frac{TP}{TP + FP} = \frac{2}{3} = 0.66$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{3}{5} = 0.6$$

ملاحظه می‌کنیم به علت کم بودن داده آموزشی دقت کم ۶۰٪ است.

۳) حین این طبقه بندی هیچ تنظیم یا Hyper Parameter نداریم که بتواند با تغییر آن از overfitting جلوگیری کرد.

مراجعه برای رفع این مشکل :

① regularization : اعمال برخی محاسبات یا مدل ساده تر یا محدود کردن لایه های درخت .

② Pruning : پس از ایجاد درخت ساختار آن را کم کنیم تا آنقدری که در طبقه بندی کردن تأثیر نداشته باشد، هرگز نمی شود

این کار باعث کوچک شدن درخت می شود.