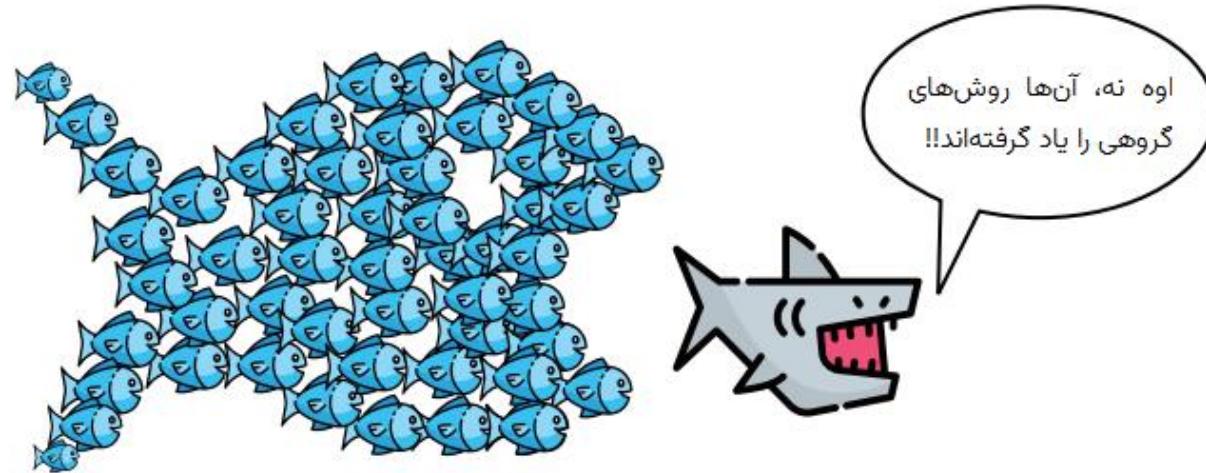


یادگیری گروهی - Adaboost



مثال

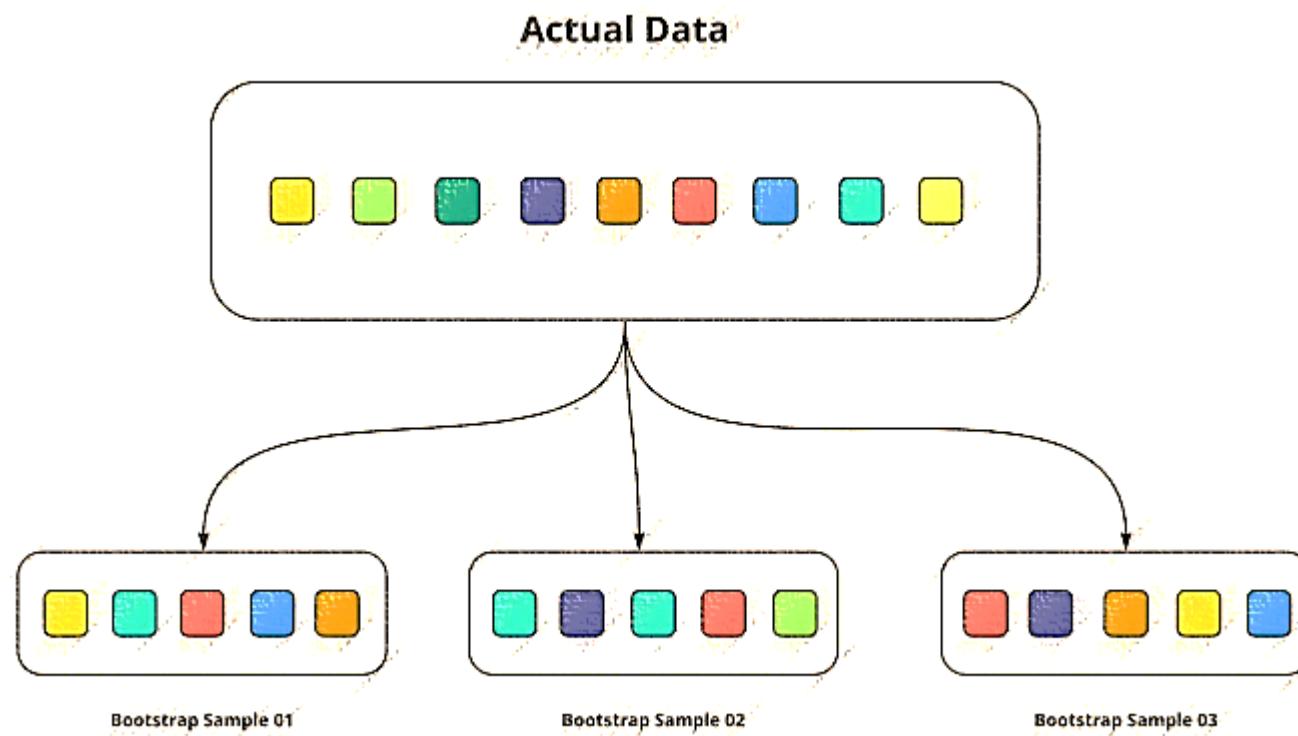
- ✓ شرکت در امتحانی با ۱۰۰ سؤال درست/غلط در موضوعات مختلف از جمله ریاضی، جغرافیا، علوم، تاریخ و موسیقی
- ✓ اجازه تماس با پنج دوست خود - آدریانا، باب، کارلوس، دانا و امیلی
- ✓ عدم نداشتن فرصت کافی جهت پاسخدهی به تمامی ۱۰۰ سؤال به دلیل اشتغال تمام وقت

دو تکنیک ممکن:

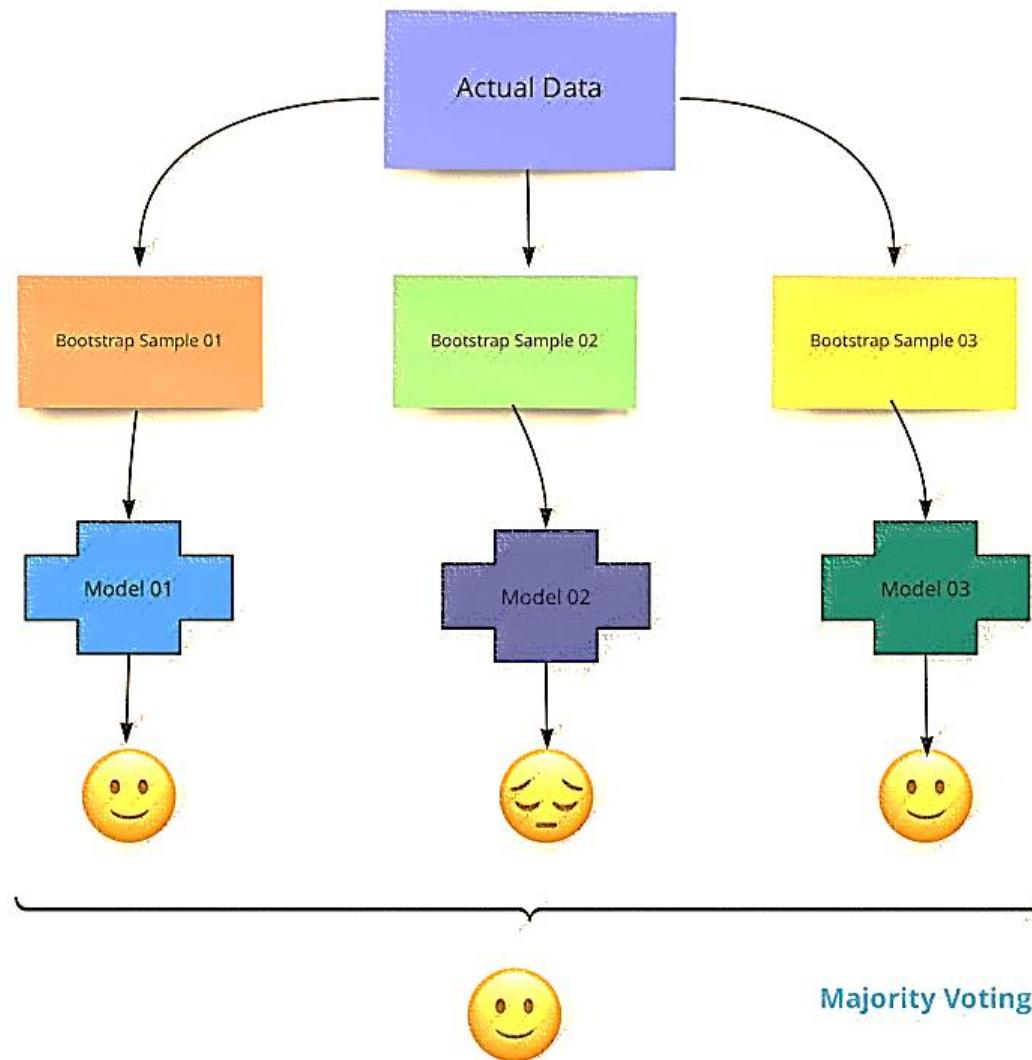
تکنیک ۱: برای هر یک از دوستان، چندین سؤال تصادفی انتخاب کنید و از آنها بخواهید که به آنها پاسخ دهند (مطمئن شوید که هر سؤال حداقل توسط یکی از دوستان ما پاسخ داده شود). پس از دریافت پاسخها، با انتخاب گزینه‌ای که در بین کسانی که به آن سؤال پاسخ داده‌اند، بیشترین محبوبیت را داشت، به آزمون پاسخ دهید. به عنوان مثال، اگر دو نفر از دوستان ما در مورد سؤال ۱ "درست" و یکی "نادرست" پاسخ داد، ما به سؤال ۱ به عنوان "درست" پاسخ می‌دهیم (در صورت وجود مساوی، می‌توانیم یکی از پاسخ‌های برنده را به طور تصادفی انتخاب کنیم).

تکنیک ۲: امتحان را به آدریانا می‌دهیم و از او می‌خواهیم که فقط به سؤالاتی که در مورد آنها مطمئن است پاسخ دهد. ما فرض می‌کنیم که آن پاسخ‌ها خوب هستند و آنها را از آزمون حذف می‌کنیم. اکنون سؤالات باقی‌مانده را با همان دستورالعمل به باب می‌دهیم. ما به این روش ادامه می‌دهیم تا آن را به هر پنج دوست منتقل کنیم.

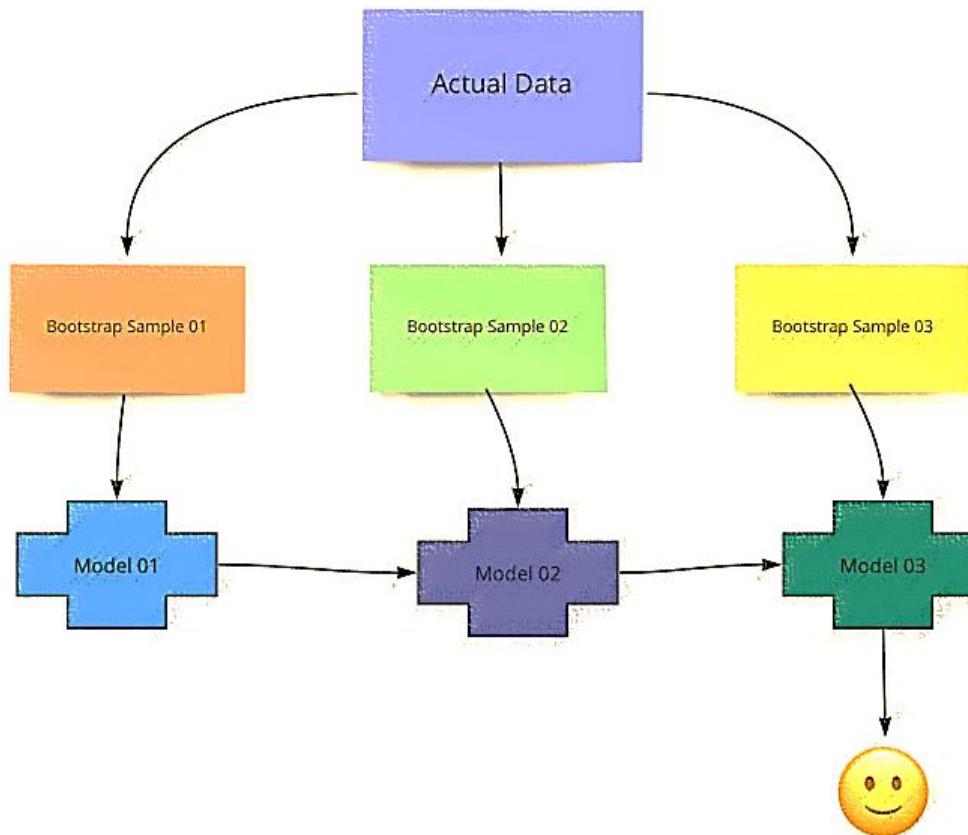
نحوه تقسيم داده



Bagging Ensemble Method



Boosting Ensemble Method

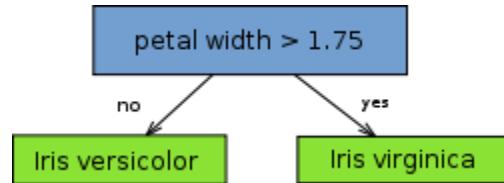


Decision Stump

استامپ تصمیم یک مدل یادگیری ماشینی است که از درخت تصمیم یک سطحی تشکیل شده است. یعنی یک درخت تصمیم با یک گره داخلی (ریشه) است که بلافاصله به گره های پایانی (برگ های آن) متصل می شود. یک استامپ تصمیم بر اساس مقدار تنها یک ویژگی ورودی پیش بینی می کند. گاهی به آنها ۱-قواعد نیز گفته می شود.

$$h(x; j, \theta) \begin{cases} +1 & x_j > \theta \\ -1 & \text{else} \end{cases}$$

نمونه ای از یک استامپ تصمیم که بین دو تا از سه کلاس مجموعه داده گل زنبق تمایز قائل می شود: **Iris virginica** و **Iris versicolor**



مثال

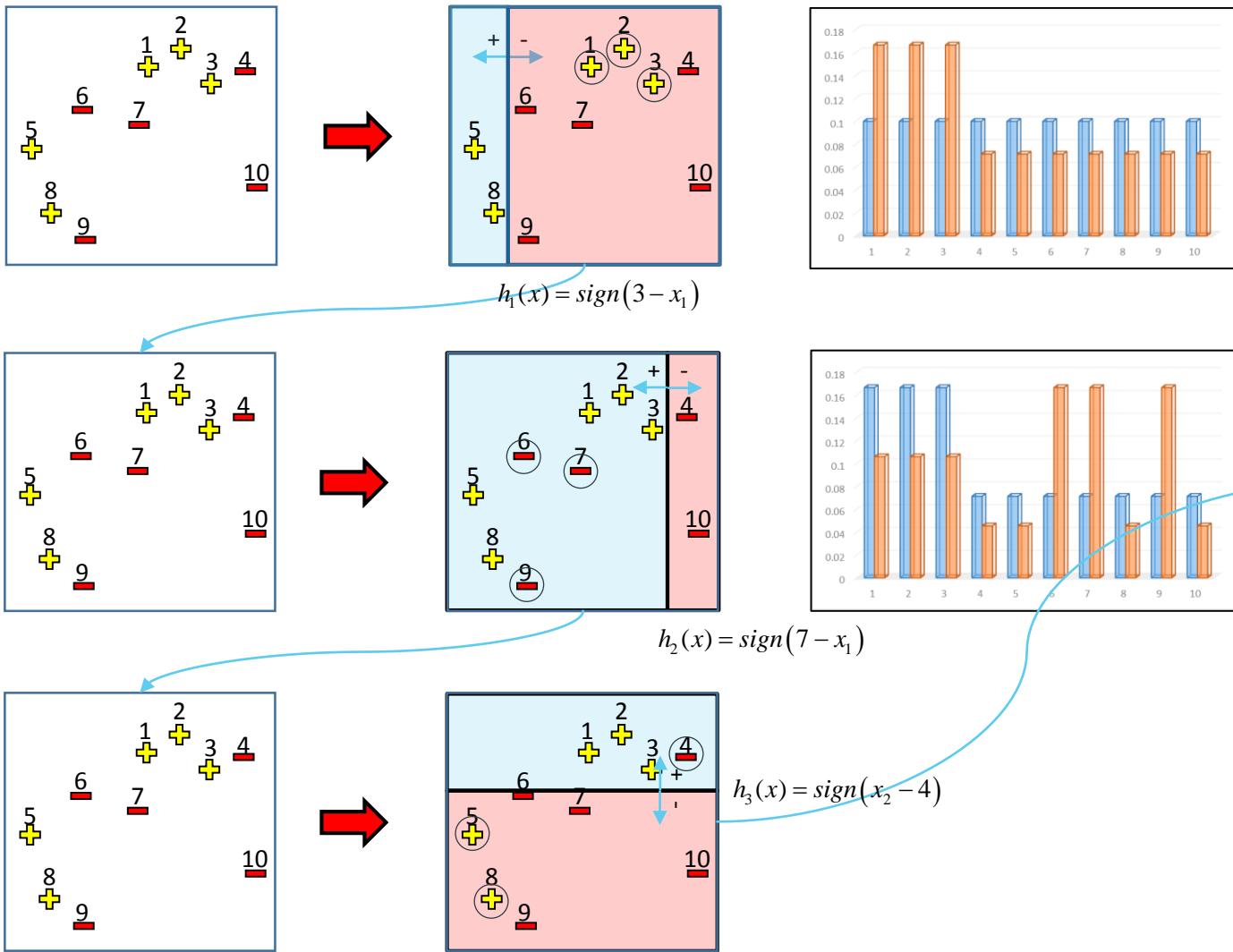
سه دوست داریم: ترزا راست‌گو، اومبرت غیرقابل‌پیش‌بینی، و لنی دروغ‌گو

ترزا راست‌گو تقریباً همیشه حقیقت را می‌گوید

اومبرت غیرقابل‌پیش‌بینی حقیقت را تقریباً نیمی از زمان می‌گوید و نیمه دیگر پنهان می‌کند

لنی دروغ‌گو تقریباً همیشه دروغ می‌گوید

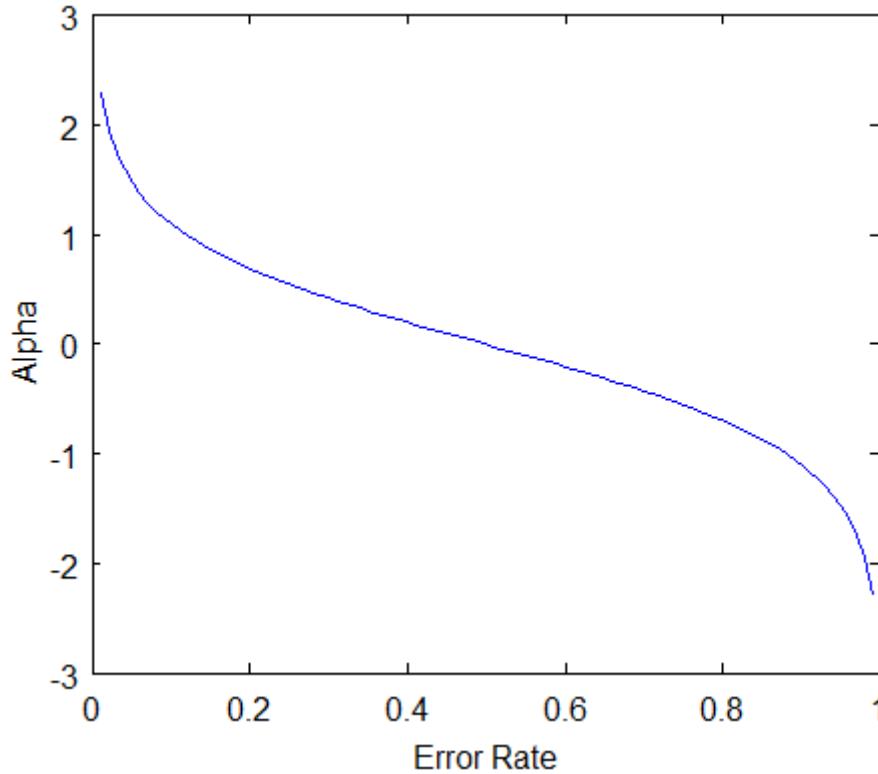
- ترزا راست‌گو نمره مثبت بالا
- لنی دروغ‌گو نمره منفی بالا
- اومبرت غیرقابل‌پیش‌بینی نمره صفر



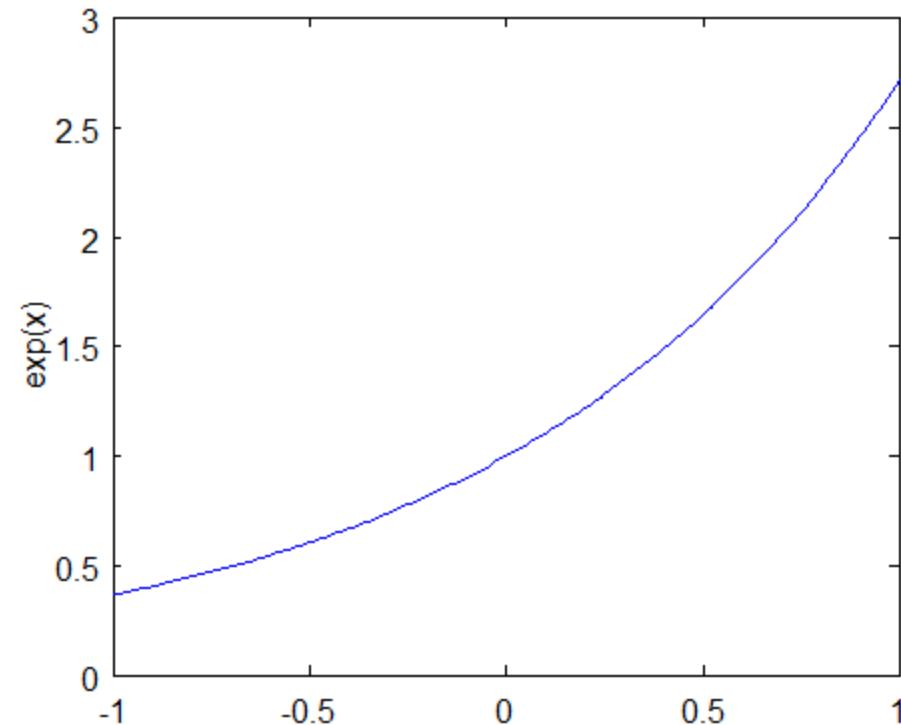
$$H_f = \text{sign}(\alpha_1 h_1(x) + \alpha_2 h_2(x) + \alpha_3 h_3(x))$$
 $= \text{sign}(0.42 \times h_1(x) + 0.6496 \times h_2(x) + 0.9223 \times h_3(x))$
 $X = \begin{cases} x_1 = 5 \\ x_2 = 2 \end{cases}$
 $H_f(X) = \text{sign}((0.42 \times -1) + (0.6496 \times 1) + (0.9223 \times -1)) = -1$

میزان خطای طبقه‌بند

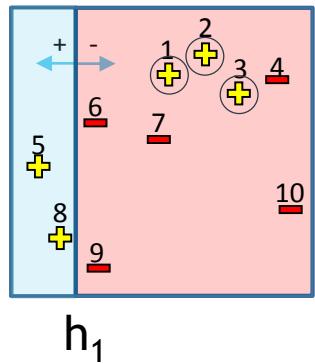
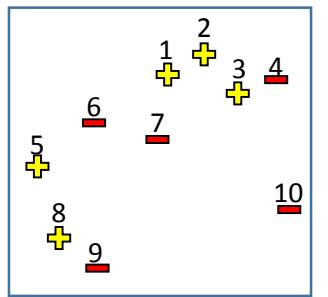
$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right)$$



درک چگونگی رفتار عبارت نمایی



10



$$J_1 = 0.1 + 0.1 + 0.1 = 0.3$$

$$\varepsilon_1 = \frac{J_1}{\sum_{i=1}^{10} Wn_{t,i}} = \frac{0.3}{1.0} < 0.5$$

$$\alpha_1 = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_1}{\varepsilon_1} \right) = \frac{1}{2} \ln \left(\frac{1 - 0.3}{0.3} \right) = 0.42$$

مجموع وزن داده‌های به اشتباه دسته‌بندی شده
خطای مناسب

ضریب دسته‌بند h_1

محاسبه وزن‌های جدید:

$$0.1 \times e^{+0.42} = 0.1528$$

$$W_{t+1,i} = W_{t,i} \times e^{+\alpha_1}, i \in \{1, 2, 3\}$$

دسته‌بندی نادرست

$$0.1 \times e^{-0.42} = 0.0655$$

$$W_{t+1,j} = W_{t,j} \times e^{-\alpha_1}, j \in \{4, 5, 6, 7, 8, 9, 10\}$$

دسته‌بندی درست

$$Z_1 = \sum_{i=1}^{10} Wun_{t+1,i} = 3 \times 0.1528 + 7 \times 0.0655 = 0.9169$$

$$Wn_{t+1,j} = \frac{Wun_{t+1,j}}{Z_1}, j = 1 \dots 10$$

نرمال کردن وزن‌ها

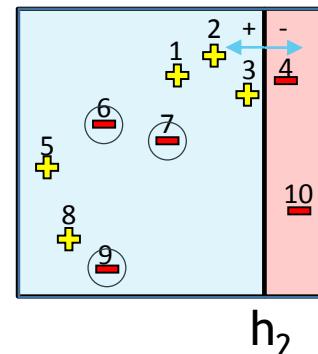
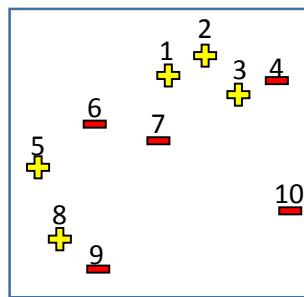
$$Wn_{t+1,1..3} = \frac{0.1528}{0.9169} = 0.1666 \quad Wn_{t+1,4..10} = \frac{0.0655}{0.9169} = 0.0714$$

1	2	3	4	5	6	7	8	9	10
0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
0.1528	0.1528	0.1528	0.0655	0.0655	0.0655	0.0655	0.0655	0.0655	0.0655
0.1666	0.1666	0.1666	0.0714	0.0714	0.0714	0.0714	0.0714	0.0714	0.0714

وزن‌های اولیه

وزن‌های جدید

نرمال کردن وزن‌ها



$$J_2 = 0.0714 + 0.0714 + 0.0714 = 0.2142$$

$$\varepsilon_2 = \frac{J_2}{\sum_{i=1}^{10} Wn_{t,i}} = \frac{0.2142}{0.9996} < 0.5$$

$$\alpha_2 = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_2}{\varepsilon_2} \right) = \frac{1}{2} \ln \left(\frac{1 - 0.2143}{0.2143} \right) = 0.6496 \quad h_2$$

مجموع وزن داده‌های به اشتباه دسته‌بندی شده

خطای مناسب

$$0.0714 \times e^{+0.6496} = 0.1367$$

دسته‌بندی نادرست

$$0.0714 \times e^{-0.6496} = 0.0373$$

دسته‌بندی درست

$$Z_2 = \sum_{i=1}^{10} Wn_{t+1,i} = 3 \times 0.0870 + 4 \times 0.0373 + 3 \times 0.1367 = 0.8203$$

$$Wn_{t+1,j} = \frac{Wun_{t+1,j}}{Z_2}, j = 1 \dots 10$$

محاسبه وزن‌های جدید:

$$Wn_{t+1,1..3} = \frac{0.0870}{0.8203} = 0.1061 \quad Wn_{t+1,4..8,10} = \frac{0.0373}{0.8203} = 0.0455 \quad Wn_{t+1,6,7,9} = \frac{0.1367}{0.8203} = 0.1666$$

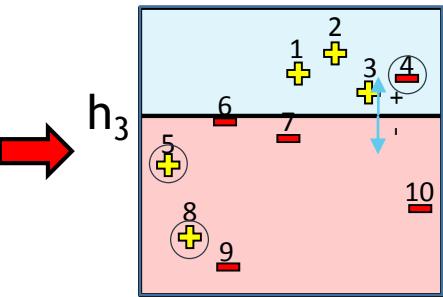
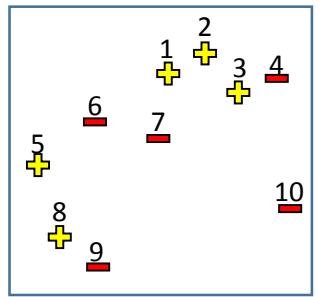
نرمال کردن وزن‌ها

1	2	3	4	5	6	7	8	9	10
0.1666	0.1666	0.1666	0.0714	0.0714	<u>0.0714</u>	<u>0.0714</u>	0.0714	<u>0.0714</u>	0.0714
0.0870	0.0870	0.0870	0.0373	0.0373	0.1367	0.1367	0.0373	0.1367	0.0373
0.1061	0.1061	0.1061	0.0455	0.0455	0.1666	0.1666	0.0455	0.1666	0.0455

وزن‌های اولیه

وزن‌های جدید

نرمال کردن وزن‌ها



$$J_3 = 0.0455 + 0.0455 + 0.0455 = 0.1365$$

$$\varepsilon_3 = \frac{J_3}{\sum_{i=1}^{10} Wn_{t,i}} = \frac{0.1365}{1.0001} < 0.5$$

$$\alpha_3 = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_3}{\varepsilon_3} \right) = \frac{1}{2} \ln \left(\frac{1 - 0.1365}{0.1365} \right) = 0.9223 \quad h_3$$

مجموع وزن داده‌های به اشتباہ دسته‌بندی شده

خطای مناسب

ضریب دسته‌بند

محاسبه وزن‌های جدید:

$$0.0455 \times e^{+0.9223} = 0.1144$$

دسته‌بندی نادرست

$$W_{t+1,i} = W_{t,i} \times e^{+\alpha_3}, i \in \{4, 5, 8\}$$

$$0.1061 \times e^{-0.9223} = 0.0422$$

دسته‌بندی درست

$$0.1666 \times e^{-0.9223} = 0.0662$$

$$W_{t+1,j} = W_{t,j} \times e^{-\alpha_3}, j \in \{1, 2, 3, 6, 7, 9, 10\}$$

$$0.0455 \times e^{-0.9223} = 0.0181$$

1	2	3	4	5	6	7	8	9	10
0.1061	0.1061	0.1061	<u>0.0455</u>	<u>0.0455</u>	0.1666	0.1666	<u>0.0455</u>	0.1666	0.0455
0.0422	0.0422	0.0422	0.1144	0.1144	0.0662	0.0662	0.1144	0.0662	0.0181

وزن‌های اولیه

وزن‌های جدید

الgoritم Adaboost

Given: $(x_1; y_1), \dots, (x_m; y_m), x_i \in X, y_i \in Y = \{-1, 1\}$.

Initialize $D_1(i) = 1/m$.

For $t = 1 \dots T$:

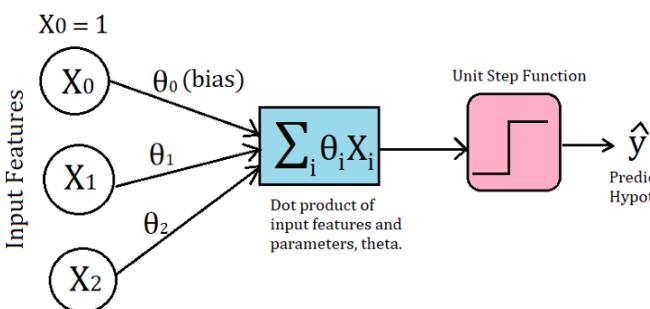
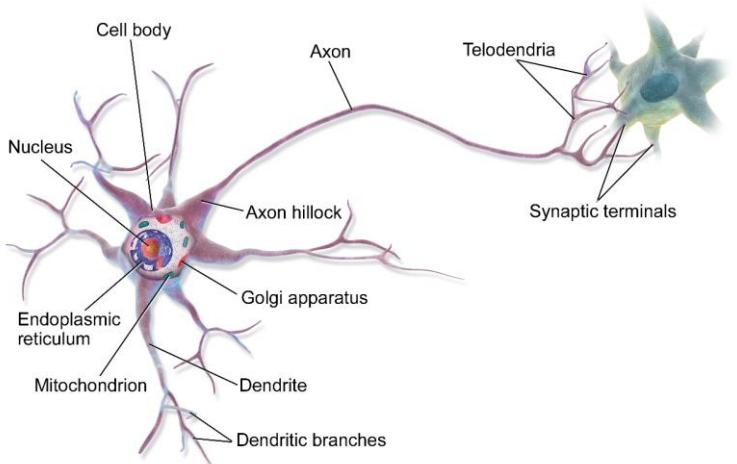
1. Train weak classifier using distribution D_t .
2. Get weak hypothesis $h_t : X \rightarrow \{-1, 1\}$ with error $\varepsilon_t = \sum_{i:h_t(x_i) \neq y_i} D_t(x_i)$
3. Choose $\alpha_t = \frac{1}{2} \log\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)$
4. Update:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \cdot = \begin{cases} e^{-\alpha_t} & \text{if instance } i \text{ is correctly classified} \\ e^{\alpha_t} & \text{if instance } i \text{ is not correctly classified} \end{cases}$$

where Z_t is a normalization factor (chosen so that $\sum_{i=1}^m D_{t+1} = 1$).

Output the final hypothesis: $H(x) = \text{sign}\left(\sum_{t=1}^T \alpha_t h_t(x)\right)$.

نگاهی گذرا به پرسپترون



مقدار تخمین زده شده
مقدار واقعی
 y
 t
 $e = t - y$
 $\theta_i \leftarrow \theta_i + \alpha \cdot e \cdot x_i$

به روزرسانی وزنها با استفاده از قانون دلتا

