

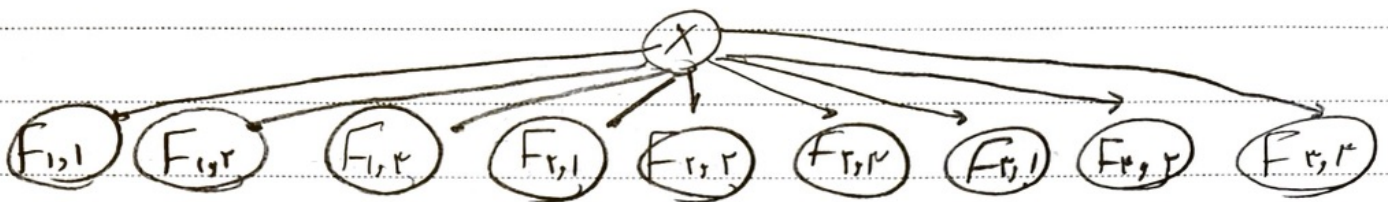
بہ نام خدا

میں بروئے جہم

سوال 1 آ در بیان گفت هر حرف ایک جدول 3×3 است که به شکل زیر تعریف میکنند:

$$F = \begin{cases} 1 & \text{if fire is on} \\ 0 & \text{if fire is off} \end{cases}$$

به شکل یک Bayes Net آن به شکل زیر است:



بیار کاربرد مدل Bayes Net 6 CPT هر node با جواب بکنیم (Smoothing = 2)

$$P(X=A) = \frac{2+2}{4+4} = 0.5$$

$$P(X=B) = \frac{2+2}{4+4} = 0.5$$

$$P(F_{on} / X=A) = \begin{bmatrix} \frac{1+2}{4+4} & \frac{2+2}{4+4} & \frac{1+2}{4+4} \\ \frac{0+2}{4+4} & \frac{2+2}{4+4} & \frac{2+2}{4+4} \\ \frac{1+2}{4+4} & \frac{0+2}{4+4} & \frac{0+2}{4+4} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

$$P(F_{on} | x=B) = \begin{bmatrix} \frac{1+r}{r+f} & \frac{1+r}{r+f} & \frac{r}{r+f} \\ \frac{r}{r+f} & \frac{1+r}{r+f} & \frac{0+r}{r+f} \\ \frac{r}{r+f} & \frac{r}{r+f} & \frac{0+r}{r+f} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{r}{V} & \frac{r}{V} & \frac{f}{V} \\ \frac{f}{V} & \frac{f}{V} & \frac{r}{V} \\ \frac{f}{V} & \frac{f}{V} & \frac{r}{V} \end{bmatrix}$$

* حال بار بار حالتی که تغییر نکند، صرفاً ضرایب $P(Y|F)$ را باید بنویسیم.

$$P(Y=A | F) \propto P(Y, F)$$

$$= P(Y=A) \times P(Y=A | f_{1,1}=0) \times P(Y=A | f_{1,2}=1) \times P(Y=A | f_{1,3}=0)$$

$$\times P(Y=A | f_{2,1}=0) \times P(Y=A | f_{2,2}=1) \times P(Y=A | f_{2,3}=1)$$

$$\times P(Y=A | f_{3,1}=1) \times P(Y=A | f_{3,2}=1) \times P(Y=A | f_{3,3}=0)$$

$$= \frac{1}{4} \times \frac{f}{V} \times \frac{f}{V} \times \frac{f}{V} \times \frac{0}{V} \times \frac{0}{V} \times \frac{0}{V} \times \frac{r}{V} \times \frac{f}{V} \times \frac{0}{V}$$

$$= \boxed{0,0029}$$

* به ترتیب $P(Y=B|F)$ را محاسبه کنیم:

$$P(Y=B|F) \propto P(Y=B, f)$$

$$= \frac{1}{2} \times \frac{4}{7} \times \frac{3}{7} \times \frac{2}{7} \times \frac{2}{7} \times \frac{3}{7} \times \frac{1}{7} \times \frac{4}{7} \times \frac{4}{7} \times \frac{1}{7}$$

$$\approx 0.0004$$

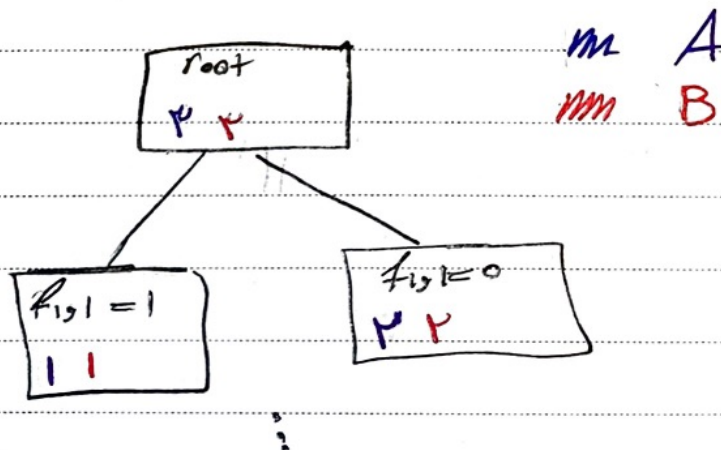
چون $P(Y=A|F) > P(Y=B|F)$ انتخاب زده شد، پس Naive Bayes

تصور دایره A انتخاب می دهد.

ب) هر کدام از یکس های موجود را با یک feature در نظر گرفته. ۵۹ داده برابر

train داریم. هر بار ۱ هر node که نیاز به split شدن داشته باشد به دو node برابر

feature بر split در شود و به دو دسته براساس A و B است:



* در نهایت هم داده ست، و در دیت traverse می کنیم تا لیست دیت برابر
تصور حدس زده شود.

سوال ۲

مقادیر بردار ابتدا تعریف می‌کنیم:

* $w^{(k)}$ و w بردار در مرحله k ام در الگوریتم* \hat{w} : w ای که \max margin را در train data set حفظ می‌کند* $\delta : \max \text{margin} = \max_w \min_i [y_i w^T x_i]$ (نویس) ** y_i : به حسب نام * x_i : مقادیر (feature) های نام

* همضرب هر مرحله از آیدیت در الگوریتم به صورت زیرات:

$$w^{(k)} = w^{(k-1)} + y_i x_i$$

* مثال انتات را شروع می‌کنیم: (فرض: $\|\hat{w}\| = 1$)* ابتدا یک گزین می‌پذیریم $w^{(k)}$ بسیار کوچک:

$$\begin{aligned} \hat{w}^T w^{(k)} &= \hat{w}^T (w^{(k-1)} + y_i x_i) \\ &= \hat{w}^T w^{(k-1)} + \hat{w}^T y_i x_i \end{aligned}$$

$$(*) \text{ (توی بار صفت)} \geq \hat{w}^T w^{(k-1)} + \delta$$

$$\delta < 1 \Rightarrow \dots \geq \hat{w}^T w^{(k-1)} + 2\delta \geq \dots$$

$$\Rightarrow \|w^{(k)}\| \geq k\delta \quad \boxed{\text{I}}$$

* حال گزین می‌پذیریم برابر $\|w^{(k)}\|$ بسیار کوچک:

$$\|w^{(k)}\|^2 = \|w^{(k-1)} + y_i x_i\|^2 = \|w^{(k-1)}\|^2 + \|y_i x_i\|^2 + 2 y_i x_i^T w^{(k-1)}$$

$$\leq \|w^{(k-1)}\|^2 + \|y_i x_i\|^2$$

(نصف کردن کلاس می‌پذیریم (ثبت شده)

$$\leq \|w^{(k-1)}\|^2 + 1$$

در گره اول به شروع

$$\leq \|w^{(k-1)}\|^2 + 2 \leq \dots \leq k$$

$$\Rightarrow \|w^{(k)}\| \leq \sqrt{k} \quad \boxed{\text{II}}$$

$$\text{I, II} \Rightarrow k\delta \leq \|w^{(k)}\| \leq \sqrt{k}$$

$$\Rightarrow k \leq \frac{1}{\delta^2}$$

که ثابت یک constant است و w در کارمده محدود، به \hat{w} همگرا شود

[نگاه از اسلایدها درسی]

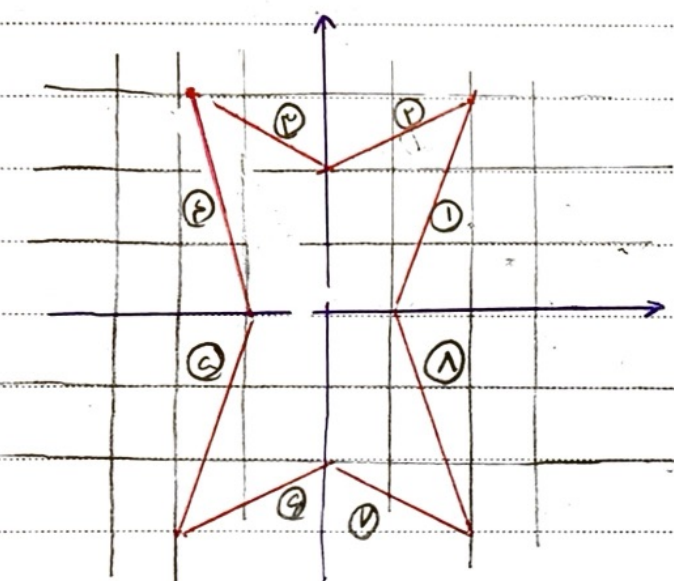
سوال ۳ [هم فکره اسیر فامرزایی]

↓ خط داریم که معادله های آن

معادله زیر مشخص شده اند.

* نکته محاسبات ترتیب تعیین کنیم
که hidden layer اول معادله های

مربعه دارند و در هر مرحله f ها دوتا
دوتا در هر hidden layer با هم ترکیب
می شوند



$$\textcircled{1} f_1(x,y) = x - x_2 + y$$

$$\textcircled{2} f_2(x,y) = x - \frac{1}{2}x - y$$

$$\textcircled{3} f_3(x,y) = x - \frac{1}{2}x - y$$

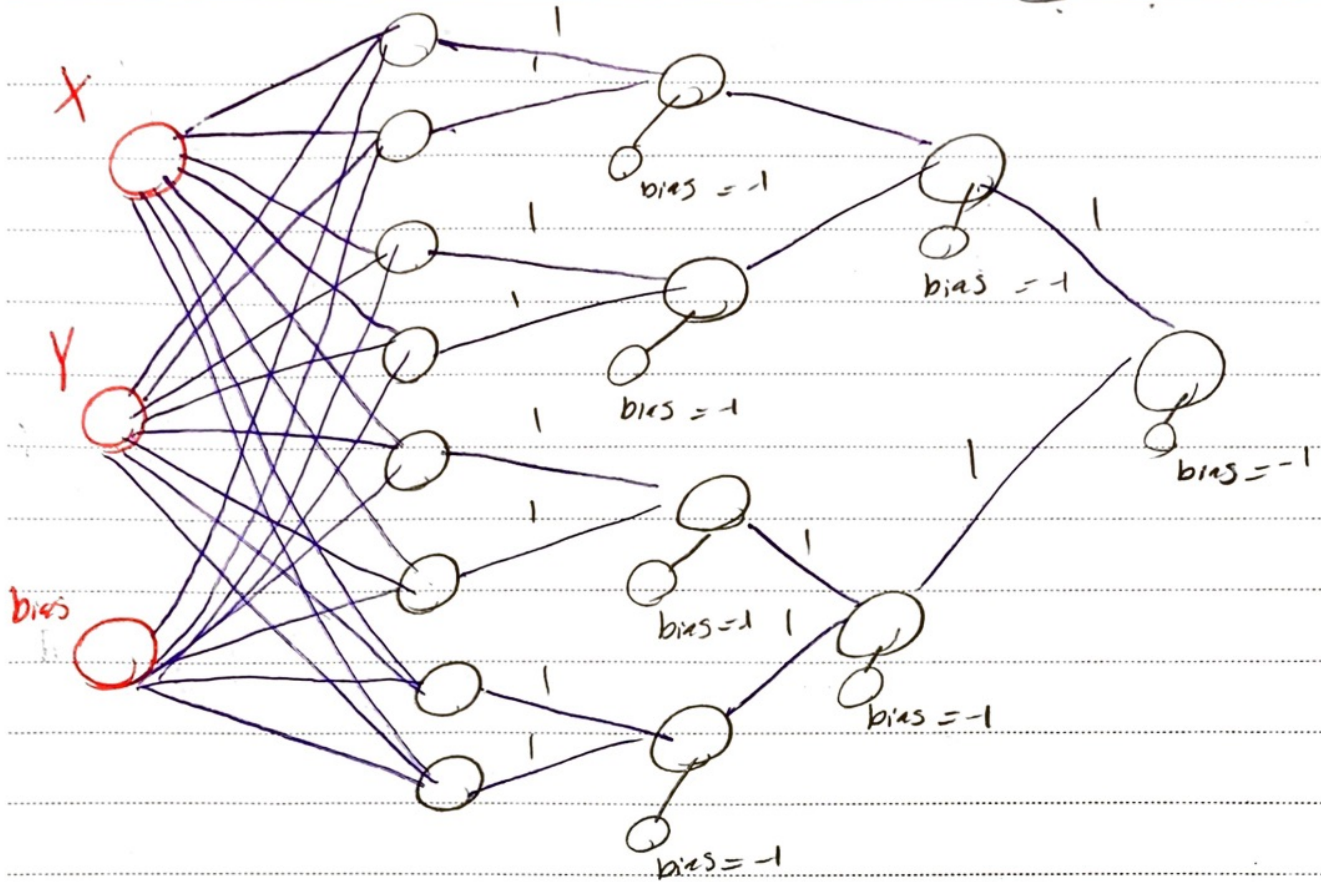
$$\textcircled{4} f_4(x,y) = x + x_2 + y$$

$$\textcircled{5} f_5(x,y) = x + x_2 - y$$

$$\textcircled{6} f_6(x,y) = x - \frac{1}{2}x + y$$

$$\textcircled{7} f_7(x,y) = x + \frac{1}{2}x + y$$

$$\textcircled{8} f_8(x,y) = x - x_2 - y$$



● وزن های bias ها در hidden layer اول، همان ضرایب bias

و بیت ها در h_1 هستند.

● وزن های bias ها و bias ها به یکدیگر وابسته است

● دلیل این موضوع این است که در هر مرحله activation function

میگیرد در هر مرحله (خوب node اگرست باشد میره از 0 تا 1 است)

اگر خوب هر دو node است باشند. حاصل جمع آن از 0 تا 1 شود

و اگر منهای از آن ها منفی باشد، از آن کم کنیم (میگردد بیشتر)

0-1 کمترین شود. □