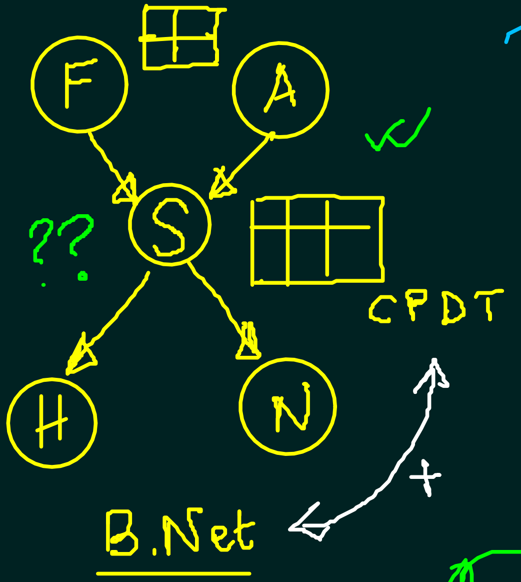


$$\overbrace{P(f, a, s, h, n)} = P(f) \cdot P(a) \cdot P(s|f,a) \cdot P(h|s) \cdot P(n|s)$$

F = f   A = a

JPDT

??

Learning

B. Net Structure ⊕ Data Set

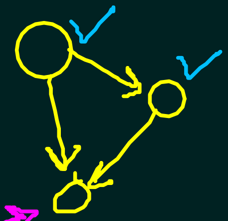
$\boxed{\text{JPDT}}$

F	A	S	H	N	Pr
1	0	1	1	1	0.4
⋮					⋮
1	0	?	1	?	

#n

Partially   Fully Obs.

CPDT $\boxed{\times}$



① B. Net Known ⊕ Fully Observed

↳ ?? Learn CPDT

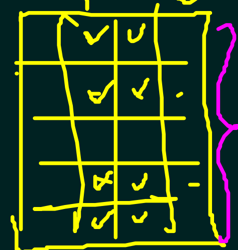
JPDT

$$\boxed{\theta(s=1 \mid F=i, A=j)}$$

$P(s|f,a)$

prob est

$$\hat{\theta}_{s|ij} = \frac{\sum_k \delta(f_k=i, a_k=j, s_k=1)}{\sum_k \delta(f_k=i, a_k=j)} \leftarrow \text{MLE}$$



$$P[\text{Data} \mid \theta] = \prod_k P(f_k, a_k, s_k, h_k, n_k)$$

$$\frac{\partial}{\partial \theta_{s|ij}} \left[\log(P(\text{Data}|\theta))\right] = \frac{\partial}{\partial \theta} \sum_k \cancel{\log P(f_k)} + \cancel{\log P(a_k)} + \underbrace{\log P(s_k \mid f_k, a_k)} + \cancel{\log P(h_k|s_k)} + \cancel{\log P(n_k|s_k)}$$

$$= \frac{\partial}{\partial \theta_{s|ij}} \left[\log P(s_k \mid f_k, a_k)\right]$$

$\boxed{\text{MLE}}$

$$\hat{\theta} = \underset{\theta}{\arg\max} \left(P(\text{Data}|\theta)\right)$$

$\underbrace{\text{T H H T H H T}}$ α_1 H α_2 T

$$\underbrace{\theta^{\alpha_1}(1-\theta)^{\alpha_2}} \quad \text{max}$$

$$\hat{\theta} = \frac{\alpha_1}{\alpha_1 + \alpha_2}$$

② Bayes Net Given ⊕ Partially Observed JPDT :

$$\hat{\Theta} = \underset{\theta}{\arg\max} \prod_k P(f_k, a_k, \cancel{s_k}, h_k, n_k)$$

PO

Itr 0.5, 0.1, 0.4

→ ① Expectation $E_{Z|X,\theta''}$

→ ② $\hat{\theta}'' \leftarrow \underset{\theta'}{\arg\max}\ E_{Z|X,\theta} \underbrace{\log\left[P(X,Z|\theta)\right]}$

exp ← log likelihood

$X = \{F, A, H, N\}$

$Z = \{S\}$

$Z = \{Y\}$

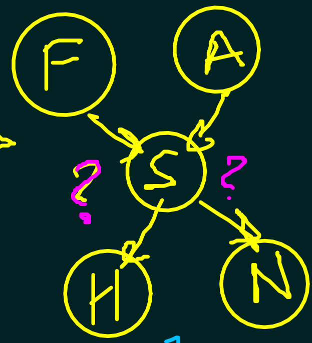
100% Y=1

0% Y=0

* E(Y=1)
* E[Y=0]

x_1	x_2	x_3	y
1	0	1	1
1	1	0	0
1	1	1	1
0	0	1	1 / 0
1	0	0	?

5



↳ Expectation Maximization (EM) algo {

① $E\left[\Rightarrow \mathrm{Prob}(Z|X,\theta)\right]$

② $\underset{\theta'}{\arg\max}\ Q[\theta'|\theta]$

$E_{Z|X,\theta} \log\left[P[X,Z|\theta]\right]$

$\theta(s|i,j)$

$$E_{s_k=1} = \frac{P(s_k=1 \mid f_k a_k h_k, n_k, \theta)}{P(s_k=1 \mid f_k a_k h_k n_k, \theta) + P(s_k=0 \mid f_k \ldots n_k \theta)}$$

Itr 1

$$= \frac{\theta(f_k)\cdot\theta(a_k)\cdot\theta(s_k=1|f_k a_k)\cdot\theta(h_k|s_k)\cdot\theta(n_k|s_k)}{\boxed{} + \text{---}\ s_k=0\ \text{------}}$$

$E[S_k]$

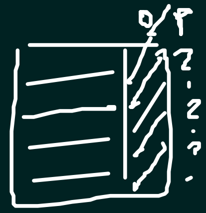
② $\theta^{(1)}_{s|ij} \leftarrow \dfrac{\sum_k \delta(f_k=i,\ a_k=j)\ E[s_k]}{\sum_k \delta(f_k=i,\ a_k=j)}$

Known $s_k = 1 \Rightarrow 1.0$

others $s_k : ? \Rightarrow$ Prev $E[s_k]$

Itr2 ① $\theta_{s|ij} \longrightarrow$ ② $\theta^{(2)}_{s|ij} \leftarrow$ - - - - - - Converge (local max.)

[θ initiation]

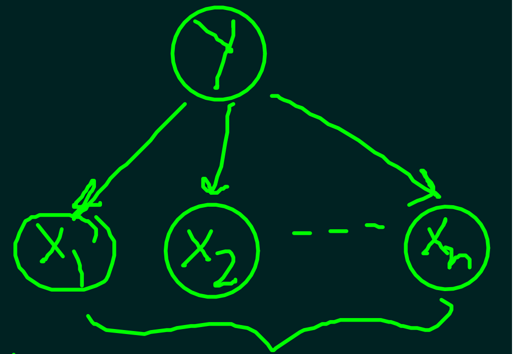


Bayes Classifier

① $E_{P(y|x_1 \ldots x_N)}[y_k = 1] = \dfrac{P[y_k=1]\ \prod_k P(x_i(k) \mid y_k=1)}{\sum_{j=0}^{1} P(y_k=j)\ \prod_k \ldots}$

Cnd.

Y	$x_1 \ldots x_n$
1	
0	
?	
?	

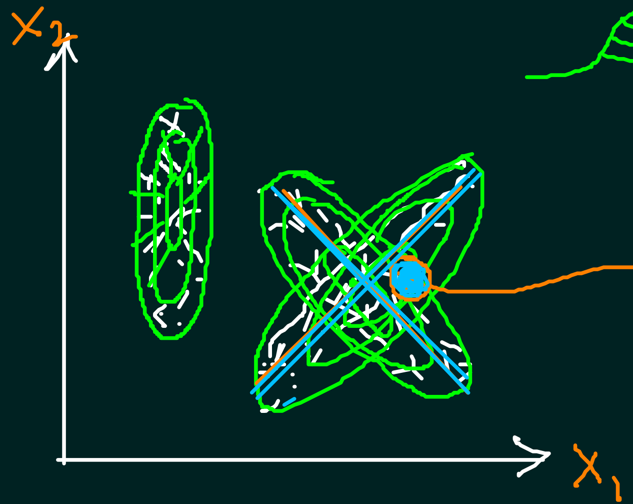


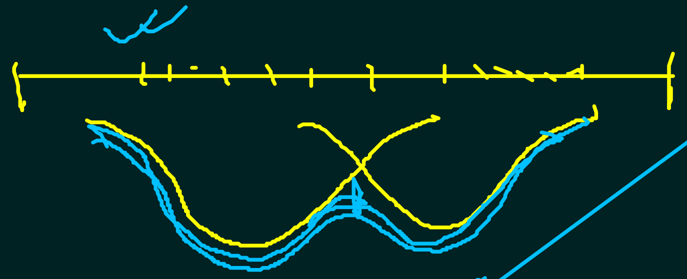
② $\theta_{x_i=j \mid y=m} = \dfrac{\sum_k P(y_k=m \mid x_1(k) \ldots x_N(k))\ \delta(x_i(k)=j)}{\sum_k P(y_k=m \mid x_1(k) \ldots x_n(k))}$

③ → B Net ?? Not Given + Fully obs Data } Chow-Lin Algo (KL Div)

↳ (Tree-Structure BN)

④ → (B Net Not Given ⊕ Partially Obs. of Data }





Mixture of Gaussian Models

$$\boxed{P(x_1, \ldots x_N)} = \underset{=?}{\sum_i} P(Z=i)\, P(x_1 \ldots x_N \mid Z)$$

1

2

$X \to$ Obs

$Z \to$ unObs

E-step: $P(Z(n) \mid X(n), \theta)$

i j k

$X(n) = \langle x_1(n) \ldots x_N(n) \rangle$

Assume

① $P(x \mid Z=j) = \prod_i N(x_i \mid \mu_{ji}, \sigma_{ji})$

$x_1 \ldots x_N$ are cond Ind.

② $\sigma_{ji} \sim \sigma$ (known)

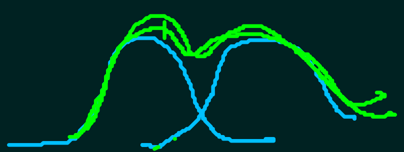
$$\boxed{P(x)} = \sum_{j=0}^{1} P(Z=j \mid \pi) \prod_i N(x_i \mid \mu_{ji}, \sigma)$$

known

Prior $(Z=j)$

Unknown

$$\theta = \langle \mu_{ji}, \pi \rangle$$



$$Q(\theta' | \theta) = E_{Z|x,\theta} \log \big[\underbrace{P(x, Z | \theta)}_{P(X|Z,\theta) \cdot P(Z|\theta)} \big] \leftarrow \underline{\text{M-step}}$$

$$\theta = \langle \mu_{ij}, \pi \rangle$$

$$= E_{Z|\theta,x} \Big[\cancel{\log \big[P(X|Z,\theta) \big]} + \log \big(P(Z|\theta) \big) \Big]$$

π

$$\frac{\partial}{\partial \pi} = 0 \longrightarrow \frac{\partial}{\partial \pi} E_{Z|\theta,x} \Big[\underbrace{\log P(Z|\theta)}_{} \Big] = 0$$

0 1 1 0 . . .

$$= \frac{\partial}{\partial \pi} \Big[\log \big(\pi^{\sum Z(n)} (1-\pi)^{\sum (1-Z(n))} \big) \Big]$$

$$\frac{\partial}{\partial \mu_{ji}} = 0$$

$$= - - - - \cdot = - - - - = \frac{1}{N} \sum_{n=1}^{N} E[Z(n)]$$

$$P(X) \leftarrow [Z_i, Z_j, Z_k, \ldots]$$

$X_1 \ldots X_n$ | $Z_1 \ldots Z_m$ $\rightarrow E[Z_i]$

1 0 0 1 - - 0 | ? 1 0 0 0 0 ? . . . 0 0 1

$\boxed{\text{Exp-Step}} \rightarrow \hat{\theta} = \langle \mu_{ji}, \pi \rangle$

$\boxed{\text{Max-Step}}$