Want probabilities, have CPTs. How do we obtain other ones?

why? less computationally intensive

	Exact	Approximate
Bayes Nets	Variable Elimination	Sampling
HMMs	Forward alg.	Particle filtering

Particle Filtering

Want to estimate probabilities of your state at time t based on evidence. Same as forward algorithm!

in expectation it's the same as forward algorithm

1) For each particle, look at transition probs., sample, and update particle position — Time Elapse Update

2) Re-weigh particles (with evidence at time t) — Observation Update

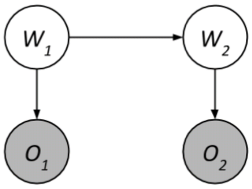
"Down weigh particles inconsistent with evidence"

3) Resample particles (so we don't have to track weighted samples)

CS188 Summer 2018 Section 10: HMMs and Naive Bayes

1 Particle Filtering

Let's use Particle Filtering to estimate the distribution of $P(W_2|O_1 = A, O_2 = B)$. Here's the HMM again:



W_1	$P(W_1)$
0	0.3
1	0.7

W_t	W_{t+1}	$P(W_{t+1}|W_t)$
0	0	0.4
0	1	0.6
1	0	0.8
1	1	0.2

W_t	O_t	$P(O_t|W_t)$
0	A	0.9
0	B	0.1
1	A	0.5
1	B	0.5

We start with two particles representing our distribution for W_1.
$P_1 : W_1 = 0$
$P_2 : W_1 = 1$
Use the following random numbers to run particle filtering:

[0.22, 0.05, 0.33, 0.20, 0.84, 0.54, 0.79, 0.66, 0.14, 0.96]

1. **Observe**: Compute the weight of the two particles after evidence $O_1 = A$.

$w(P_1) = P(O_1 = A \mid W_1 = 0) = 0.9$

$w(P_2) = P(O_1 = A \mid W_1 = 1) = 0.5$

2. **Resample**: Using the random numbers, resample P_1 and P_2 based on the weights.

Want to use the weights above to sample new particle positions

0.22
0.05

Ranges:

$W_1 = 0 \sim \frac{0.9}{0.5+0.9} = 0.64 \Rightarrow [0, 0.64)$

$W_2 = 1 \sim \frac{0.5}{0.5+0.9} = 0.36 \Rightarrow [0.64, 1)$

Resampling

$P_1 = 0$ (since $0.22 < 0.64$)

$P_2 = 0$ (since $0.05 < 0.64$)

3. **Elapse Time**: Now let's compute the elapse time particle update. Sample P_1 and P_2 from applying the time update.

$P_1 = 0$ (since $0.33 < 0.4$)

$P_2 = 0$ (since $0.20 < 0.4$)

W_t	W_{t+1}	$P(W_{t+1}\|W_t)$
0	0	0.4
0	1	0.6
~~1~~	~~0~~	~~0.8~~
~~1~~	~~1~~	~~0.2~~

0.33 0.20

what if 0.70

4. **Observe**: Compute the weight of the two particles after evidence $O_2 = B$.

$w(P_1) = P(O_2 = B \mid W_1 = 0) = 0.1$

$w(P_2) = P(O_2 = B \mid W_1 = 0) = 0.1$

W_t	O_t	$P(O_t\|W_t)$
0	A	0.9
0	B	0.1
1	A	0.5
1	B	0.5

5. **Resample**: Using the random numbers, resample P_1 and P_2 based on the weights.

Intervals?

0.84, 0.54,

Just $W_1 = 0$! $[0, 1]$

$\Rightarrow$ $P_1 = 0$

$P_2 = 0$

6. What is our estimated distribution for $P(W_2 | O_1 = A, O_2 = B)$?

$$P(W_2 = 0 \mid O_1 = A, O_2 = B) = \frac{2}{2} = 1$$

$$P(W_2 = 1 \mid O_1 = A, O_2 = B) = \frac{0}{2} = 0$$

Naive Bayes

Objective → Classification (labeling data based on training data)

Preprocessing → Based on features

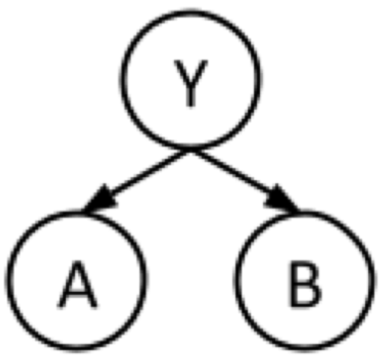
Model → strong independence assumptions (bayes net)

Method → MLE (Maximum likelihood Estimation

2 Naive Bayes

In this question, we will train a Naive Bayes classifier to predict class labels Y as a function of input features A and B. Y, A, and B are all binary variables, with domains 0 and 1. We are given 10 training points from which we will estimate our distribution.

A	1	1	1	1	0	1	0	1	1	1
B	1	0	0	1	1	1	1	0	1	1
Y	1	1	0	0	0	1	1	0	0	0



1. What are the maximum likelihood estimates for the tables $P(Y)$, $P(A|Y)$, and $P(B|Y)$?

$$\mathcal{L}(x_1, \dots x_n) = \prod_i P(x_i)$$

$$= \prod_i P(A=a_i, B=b_i, Y=y_i)$$

$$= \prod_i P(A=a_i | Y=y_i)\, P(B=b_i | Y=y_i)\, P(Y=y_i)$$

$\theta = P(F=f \mid Z=z')$ (in this case $\theta_1 = P(Y=0)$, $\theta_2 = P(A=1 \mid Y=0)$, ...)

$$\theta = \underset{\theta}{\arg\max}\ \mathcal{L}_\theta(x_1, \dots, x_n) = \frac{1}{N_{z'}} \sum_i^{N_{z'}} f_i \qquad P_{MLE}(x) = \frac{\text{count}(x)}{N}$$

A	1	1	1	1	0	1	0	1	1	1
B	1	0	0	1	1	1	1	0	1	1
Y	1	1	0	0	0	1	1	0	0	0

Y	$P(Y)$
0	6/10
1	4/10

A	Y	$P(A \mid Y)$
0	0	1/6
1	0	5/6
0	1	1/4
1	1	3/4

B	Y	$P(B \mid Y)$
0	0	2/6
1	0	4/6
0	1	1/4
1	1	3/4

2. Consider a new data point ($A = 1$, $B = 1$). What label would this classifier assign to this sample?

What label CAN we assign?

What tools do we have to decide?

$$P(Y=0, A=1, B=1) = \frac{6}{10} \cdot \frac{5}{6} \cdot \frac{4}{6} = \frac{1}{3}$$

$$P(Y=1, A=1, B=1) = \frac{4}{10} \cdot \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{40}$$

3. Let's use Laplace Smoothing to smooth out our distribution. Compute the new distribution for $P(A|Y)$ given $\underbrace{\text{Laplace Smoothing}}_{\text{prevent overfitting!}}$ with $k = 2$.

$$P_{LAP,k}(x) = \frac{\text{count}(x) + k}{N + k|X|}$$

$$P_{LAP,0}(x) = P_{MLE}(x)$$

$$\lim_{k \to \infty} P_{LAP,k}(x) = \frac{1}{|X|}$$

A	Y	$P(A \mid Y)$
0	0	
1	0	
0	1	
1	1	

$$\frac{1+2}{6+2\cdot 2} = \frac{3}{10}$$

$$\frac{5+2}{6+2\cdot 2} = \frac{7}{10}$$

$$\frac{1+2}{4+2\cdot 2} = \frac{3}{8}$$

$$\frac{5}{8}$$

A	1	1	1	1	0	1	0	1	1	1
B	1	0	0	1	1	1	1	0	1	1
Y	1	1	0	0	0	1	1	0	0	0