Week 2

Monday, June 22, 2020 11:58 PM

1

Probability

having possible worlds (w)
And the possibility of some possible world is p(w)

$0 <= p(w) <= 1$

$$\sum_{\omega \in \Omega} p(w) = 1$$

Unconditional probability :

Degree of belief in a proposition in the absence of any other evidence

Conditional probability :

Degree of belief in a proposition given some evidence that has already been revealed

P(a | b) --> probability of a given b

P(a | b) = $\frac{p(a \wedge b)}{p(b)}$

P(a ^ b) = P(b)p(a | b) = p(a)p(b | a)

Random variable :

A variable in probability theory with a domain of possible values it can take on

Probability distribution :

Takes a random variable and give the probability for each value in its domain

P(Var) = <-,-,-,->

Independence :

The knowledge that one event occurs does not affect the probability of the other Event

P(a ^ b) = p(b)p(a | b)
P(a | b) = p(a)
So :

p(a ^ b) = p(a)p(b)

Bayes` Rule :

P(a)p(b | a) = p(b)p(a | b)
So :

P(b | a) = $\frac{p(b)p(a \mid b)}{p(a)}$

Usage : Express conditional probability given the reverse of it

Knowing :
P(visible effect | unknown cause)
We can calculate :
P(unknown cause | visible effect)

Joint probability :

Considering the likelihood of several deferent events at the same time
P(var1 ^ var2) = P(var1, var2) = <-, -, -, ->

Given joint probability table :
We Can draw information about conditional probability
P(Var | something) = $\frac{p(var, somthing)}{p(something)}$
P(something) is a constant so :
P(Var | something) = α p(Var ^ something)

α : normalization factor

Probability Rules :

Negation :
P($\neg a$**) = 1 - p(a)**

Inclusion-Exclusion :
P(a V b) = p(a) + p(b) - P(a ^ b)

Marginalization :
P(a) = p(a ^ b) + p(a ^ $\neg b$)

P(X = x_i**) =** $\sum_j p(X = Xi, Y = yj)$

Conditioning :
P(a) = p(a | b)p(b) + p(a | $\neg b$)p($\neg b$)

P(X = x_i**) =** $\sum_j p(X = Xi \mid Y = yj)$**p(**$Y_j$**)**

2

Bayesian network :

Data structure that represent the dependencies among random variable

Design of Bayesian network:
- Directed graph
- Each node represent a random variable
- Arrow from X to Y means X is a parent of Y
- Each node X has probability distribution P(X | parents(X))

Approximate inference :

Do not know the exact probability but I have a general sense for the probability and can get better with time

Example : sampling

Takes sample by taking a value of every node

(rejection sampling) :

- Arrow from X to Y means X is a parent of Y
- Each node X has probability distribution P(X | parents(X))

Inference in probabilistic sitting :

- **Query(x)** : variable for which to compute distribution
- **Evidence variables E** : observed variables for event e
- **Hidden Variables Y** : non-evidence, non-query variable
- **Goal : p(X | e)**

EX : Inference By Enumeration :

$P(X \mid e) = \alpha\, p(X, e) = \alpha \sum_{y} p(X, e, y)$

y : ranges of values of hidden variables
α : Normalizing the result

Example : sampling

Takes sample by taking a value of every node

(rejection sampling) :

From the samples : reject the samples that does not
Match the evidence

(likelihood weighting) :

a. Start by fixing the value for the evidence variables(and sample them)
b. Sample the non-evidence variables using conditional probability in the Bayesian network
c. Weight each sample by its likelihood (The probability of all of the evidence)

Uncertainty over time :

X_t : the variable X at time t

Markov assumption :

The assumption that the current state depend on only
A finite fixed number of previous states

Markov chain :

A sequence of random variables where the distribution of each variable follow the Markov assumption

Transition Model :

How we transition from one state to next state

Sensor Models :

Translate the hidden state to an observation

Hidden Markov Model :

A Markov model for a system with hidden states
That generate some observed event

--- We need another model between
state and event --> **sensor model**(called: emission probability)

Sensor Markov assumption :

The assumption that the evidence variable depend only on the corresponding state

Task	Definition
Filtering	Given an observations from start Until now, calculate distribution for **current** state
Prediction	Given an observations from start Until now, calculate distribution for **future** state
smoothing	.... For past state
Most likely Explanation	Given an observations from start Until now, calculate most likely sequence of states

---