Topic 4

Representation and Reasoning with Uncertainty

Contents

4.1 Probabilistic methods (Bayesian)

Alternative form of Bayes Rule

- Bayes Rule has the form:

$$p(X \mid Y) = \frac{p(Y \mid X) \cdot p(X)}{p(Y)}$$

- But to calculate this, we need p(X) and p(Y)
- Often, p(Y) is not available.
- Sometimes we might not have p(Y) but have p(X|¬ Y)
  - E.g., don't know p(has_cold), but we do know p(sneeze|has_cold)
  - (a doctor who only sees sick people cannot estimate what proportion of the world have colds, but can estimate what proportion of her patients with colds are sneezing)
- An alternative version of Bayes rule can be derived.

4.1 Probabilistic methods (Bayesian)

Alternative form of Bayes Rule

- An alternative version of Bayes rule can be derived.
  1. Bayes Rule:

$$p(X \mid Y) = \frac{p(Y \mid X) \cdot p(X)}{p(Y)}$$

  2. We know: p(Y) = p(Y|X)*P(X) + p(Y| ¬X)*p(¬X)
  3. Substitute into Bayes Rule

$$p(X \mid Y) = \frac{p(Y \mid X) \cdot p(X)}{p(Y \mid X) \cdot p(X) + p(Y \mid \neg X) \cdot p(\neg X)}$$

In this form, we just need p(X), p(Y|X) and p(Y|¬X)

4.1 Probabilistic methods (Bayesian)

Example using Alternative form of Bayes Rule

- Bayes Rule Alternative Form

$$p(X \mid Y) = \frac{p(Y \mid X) \cdot p(X)}{p(Y \mid X) \cdot p(X) + p(Y \mid \neg X) \cdot p(\neg X)}$$

- Given the following data,

  p(low oil) = 0.3

  p(overheat | low oil) = 0.85

  p(overheat | ¬ low oil) = 0.2

Use Bayes alternative form to calculate

  p(low oil | overheat)

Topic 4

Representation and Reasoning with Uncertainty

4.1 Probabilistic methods (Bayesian)

Multiple Sources of Evidence

4.1 Probabilistic methods (Bayesian)

What happens when we have more than one source of evidence?

Sometimes we have multiple sources of evidence for our hypothesis,

E.g., if sneezing then has_cold (0.8)

if has_fever then has_cold (0.3)

How do we combine?

4.1 Probabilistic methods (Bayesian)

Combining evidence sources in conditional probabiliites

Method I: Treat E as E_1 & E_2 & E_3 ....

$$p(H \mid E) = p(E \mid H) * p(H) / p(E)$$

$$\Rightarrow \quad p(H \mid E_1 \,\&\, E_2 \,\&\, E_3 \ldots) = \frac{p(E_1 \,\&\, E_2 \,\&\, E_3 \mid H) * p(H)}{p(E_1 \,\&\, E_2 \,\&\, E_3)}$$

$$\Rightarrow \quad p(H \mid E_1 \,\&\, E_2 \,\&\, E_3 \ldots) = \frac{p(E_1 \,\&\, E_2 \,\&\, E_3 \mid H) * p(H)}{p(E_1 \,\&\, E_2 \,\&\, E_3 \mid H) + p(E_1 \,\&\, E_2 \,\&\, E_3 \mid \neg H)}$$

BUT: this would requires us to know 2^{n+1} distinct probabilities, e.g.

probability of sneezing/no_fever given cold
probability of sneezing/fever given cold
probability of not_sneezing/fever given cold
probability of not_sneezing/no_fever given cold

etc.

4.1 Probabilistic methods (Bayesian)

$$p(H \mid E_1, E_2, \ldots, E_n) = \frac{p(H) \cdot p(E_1, \ldots, E_n \mid H)}{p(H) \cdot p(E_1, \ldots, E_n \mid H) + p(\neg H) \cdot p(E_1, \ldots, E_n \mid \neg H)}$$

Let us consider the following problem:

p(infection, pain | caries) = 0.52	*p*(infection, pain | ¬caries) = 0.05
p(infection, ¬pain | caries) = 0.25	*p*(infection, ¬pain | ¬caries) = 0.1
p(¬infection, pain | caries) = 0.15	*p*(¬infection, pain | ¬caries) = 0.2
***p*(¬infection, ¬pain | caries) = 0.08**	***p*(¬infection,¬pain| ¬caries) =0.65**

p(caries) = 0.3, ***p*(¬caries) = 0.7**

If we observe pain and infection, what is the probability of caries?

p(caries | infection, pain) = 0.3*0.52/(0.3*0.52+0.7*0.05) = 0.82

If patient has pain but no infection:

p(caries | pain, ¬infection) = 0.3*0.15/(0.3*0.15+0.7*0.2) = 0.08

4.1 Probabilistic methods (Bayesian)

$$p(H \mid E_1, E_2, \ldots, E_n) = \frac{p(H) \cdot p(E_1, \ldots, E_n \mid H)}{p(H) \cdot p(E_1, \ldots, E_n \mid H) + p(\neg H) \cdot p(E_1, \ldots, E_n \mid \neg H)}$$

- **The problem** is that, assuming the E_i are boolean, we would need to know the value of **2^{n+1} probabilities.**
- If we have 10 possible variables, we need to measure 2048 probabilities.
- Real problems can have hundreds or thousands of variables!
- Very pretty, but theoretically intractable.

9

4.1 Probabilistic methods (Bayesian)

Alternative formula: Bayes Incremental Rule

- A means of approximating the conditional probability of a set of evidences.
- Starts with no evidence, and adds in each evidence source one at a time.
- Complexity is far lower.

Old form: $$p(H \mid E_n) = \frac{p(H) \cdot p(E_n \mid H)}{p(E_n)}$$

Incremental form: $$p(H \mid E_n, E_O) = \frac{p(H \mid E_O) \cdot p(E_n \mid H, E_O)}{p(E_n \mid E_O)}$$

10

4.1 Probabilistic methods (Bayesian)

Alternative formula: Bayes Incremental Rule

$$p(H \mid E_n, E_O) = \frac{p(H \mid E_O) \cdot p(E_n \mid H, E_O)}{p(E_n \mid E_O)}$$

- E_O can be interpreted as an event that consists of the simultaneous observation of a set of evidences E_1, E_2, E_3 ... E_{n-1}

  I.E., $E_O = E_1 \wedge E_2 \wedge E_3 \ldots \wedge E_{n-1}$
- E_n we can interpret as the observation of an additional evidence that is presented to us after having observed the set of evidences E_O.
- Therefore, we have total evidence $E = E_n \wedge E_O$
- The equation tells us **how it changes our belief** in H when we are given a new evidence E_n.

  If with the evidence E_O our belief in H *is* $p(H \mid E_O)$, when we observe some new evidence E_n, we should change our belief to $p(H \mid E_n, E_O)$ in this way

4.1 Probabilistic methods (Bayesian)

Bayes Incremental Rule: **Assuming Independence**

$$p(H \mid E_n, E_O) = \frac{p(H \mid E_O) \cdot p(E_n \mid H, E_O)}{p(E_n \mid E_O)}$$

- Assuming that each evidence source is independent:

  $p(E_n|E_O) = p(E_n)$

  $p(E_n|H,E_O) = p(E_n|H)$
- Thus:

$$\boxed{p(H|E_n|E_O) = \frac{p(H|E_O) \;.\; p(E_n|H)}{p(E_n)}}$$

4.1 Probabilistic methods (Bayesian)

$$p(H \mid E_n, E_O) = \frac{p(H \mid E_O) \cdot p(E_n \mid H)}{p(E_n)}$$

How do we use this rule?

1) Initially $E_O \leftarrow \{\}$. Thus $p(H \mid E_O) \leftarrow p(H)$
2) We are presented with the first evidence E_1. Update :
   $E_O \leftarrow \{E_1, E_O\} = \{E_1\}$
   $p(H \mid E_O) \leftarrow p(H \mid E_O) * p(E_1 \mid \mathrm{H}) / p(E_1)$
3) We are presented with the 2nd evidence E_2. Update:
   $E_O \leftarrow \{E_2, E_O\} = \{E_2, E_1\}$
   $p(H \mid E_O) \leftarrow p(H \mid E_O) * p(E_2 \mid \mathrm{H}) / p(E_2)$

…

Problem: often the evidences are not independent, and using this formula 1) will give errors, and 2) it can give absurd values > 1.

13

Incrementatl Bayes: normalised and assuming independence

$$p(H \mid E_n, E_O) = \frac{p(H \mid E_O) \cdot p(E_n \mid H)}{p(H \mid E_O) \cdot p(E_n \mid H) + p(\neg H \mid E_O) \cdot p(E_n \mid \neg H)}$$

$$p(\neg H \mid E_n, E_O) = 1 - p(H \mid E_n, E_O)$$

1) Initially $E_O \leftarrow \{\}$. $p(H \mid E_O) \leftarrow p(H)$, $p(\neg H \mid E_O) \leftarrow 1 - p(H)$
2) We are presented with the first evidence E_1. Update:
   $E_O \leftarrow \{E_1 , E_O\} = \{E_1\}$
   $p(H \mid E_O) \leftarrow p(H \mid E_O) * p(E_1 \mid \mathrm{H}) /$ …
   $p(\neg H \mid E_O) \leftarrow 1 - p(H \mid E_O)$
3) We are presented with the second evidence E_2. Update:
   $E_O \leftarrow \{E_2 , E_O\} = \{E_2 , E_1\}$
   $p(H \mid E_O) \leftarrow p(H \mid E_O) * p(E_2 \mid \mathrm{H}) /$ …
   $p(\neg H \mid E_O) \leftarrow 1 - p(H \mid E_O)$

…

The solution is still an estimate, but no longer gives absurd values

14

Example:

p (sneezing) = 0.4
p (cold) = 0.3
p (sneezing | cold) = 0.75
p(fever| ¬cold) = 0.2
p(fever| cold) = 0.7

$$p(H \mid E_n, E_O) = \frac{p(H \mid E_O) \cdot p(E_n \mid H}{p(E_n)}$$

What is the probability of having a cold if sneezing and fever?

p(cold|sneezing) = 0.56 (from prior calculation)
p(cold|sneezing,fever)= 0.56 *0.7 / (0.56*0.7 + 0.44*0.2)
=0.392/(0.392+0.088)
=0.817

$$p(H \mid E_n, E_O) = \frac{p(H \mid E_O) \cdot p(E_n \mid H)}{p(H \mid E_O) \cdot p(E_n \mid H) + p(\neg H \mid E_O) \cdot p(E_n \mid \neg H)}$$