CMPT 423/820

Assignment 1 Question 8 - Model Solution and Grading

Task

Derive an expression for $P(\mu|\mathbf{X}_1, \mathbf{X}_2)$ in terms of yesterday's posterior $P(\mu|\mathbf{X}_1)$. This expression shows how yesterday's posterior can be used as if it were a prior.

Model Solution

We'll proceed as follows:

$$\begin{aligned}
P(\mu|\mathbf{X}_1, \mathbf{X}_2) &= \frac{P(\mathbf{X}_2\mathbf{X}_1\mu)}{P(\mathbf{X}_2\mathbf{X}_1)} \\
&= \frac{P(\mathbf{X}_2|\mathbf{X}_1\mu)P(\mu\mathbf{X}_1)}{P(\mathbf{X}_2\mathbf{X}_1)} \\
&= \frac{P(\mathbf{X}_2|\mu)P(\mu\mathbf{X}_1)}{P(\mathbf{X}_2\mathbf{X}_1)} \\
&= \frac{P(\mathbf{X}_2|\mu)P(\mu|\mathbf{X}_1)P(\mathbf{X}_1)}{P(\mathbf{X}_2\mathbf{X}_1)}
\end{aligned}$$

The first line is the definition of conditional probability. The second line is the product rule. The third line is an assumption of conditional independence, that $P(\mathbf{X}_2|\mathbf{X}_1\mu) = P(\mathbf{X}_2|\mu)$. The fourth line is the product rule.

We could go one step further, using the definition of conditional probability, as follows:

$$P(\mu|\mathbf{X}_1, \mathbf{X}_2) = \frac{P(\mathbf{X}_2|\mu)P(\mu|\mathbf{X}_1)}{P(\mathbf{X}_2|\mathbf{X}_1)}$$

This gives us Bayes Rule in a conditional context.

But that's not really important here. The main thing is that we have $P(\mu|\mathbf{X}_1, \mathbf{X}_2)$ on the left side, and $P(\mu|\mathbf{X}_1)$ on the right. IN other words, we can start learning today, using the model we had from yesterday. I think that's a good thing!

A plausible derivation can use the same rules in different orders. For example, it's valid to use the assumption of conditional independence as follows: $P(\mathbf{X}_2\mathbf{X}_1|\mu) = P(\mathbf{X}_2|\mu)P(\mathbf{X}_1|\mu)$. A derivation that does this can be correct, but will probably be a lot longer!

Grading:

1. Grading should focus on the use of the basic rules of probability:

    - Bayes Rule
    - Product Rule
    - Conditional Independence
    - Definition of Conditional Probability

    No other skills are needed here. No mention of Beta distributions, or expected values are needed.
2. Deduct a mark for every incorrect use of a basic rule.
3. If the derivation is incomplete, but more or less correct, give 3 marks.
4. Deductions should not bring grade below 0!