

**Goal:**

Give a model

$p(y|x)$  : on "my understanding level  $y$  after class given condition  $x$ . "

This is supposed to help me decide whether I should attend a lecture.

- $\mathcal{X}$  be my "condition" before lecture. Each  $x \in \mathcal{X}$  consists enough data, where for  $i = 1, \dots, N$ , I can define a collection of functions

$$\{q_i : \mathcal{X} \rightarrow \mathbb{R}\}_{i=1}^N$$

- $\mathcal{Y} := \{0, 1\}$ , 0 means I basically understood nothing, and 1 means I got something out.

**Working example**

To give an explicit example: let  $N = 5$ .  $\mathcal{X} := \mathbb{R}^5$ , for  $i = 1, \dots, 5$   $q_i : \mathbb{R}^5 \rightarrow \mathbb{R}$  be simply projection on to the  $i$ th component,  $x = (x_1, \dots, x_5) \mapsto x_i$ .

$x = (x_1, \dots, x_5) \in \mathcal{X}$  encodes the following data:

- $x_1$ : mood in scale 1-10.
- $x_2$ : sleepiness in scale 1-10.
- $x_3$ : zoom-ness (that is 1 if lecture was on zoom, 0 otherwise)
- $x_4$ : my background knowledge for the coming lecture.
- $x_5$  day since first class.

The training set can be collected from attending each lecture.

**Some feature choices**

We have the following features: where

$$f(-, -) : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}^k$$

All features are binary. Let me give five most important features  $f_1, \dots, f_5$  continuing the above explicit example. First we can have features that are dependent on my mood.

$$f_{1a}(x, y) = \begin{cases} 1 & \text{if } q_1(x) \geq 5 \wedge y = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_{1b}(x, y) = \begin{cases} 1 & \text{if } q_1(x) \leq 5 \wedge y = 0 \\ 0 & \text{otherwise} \end{cases}$$

But one is really the "negation" of the other. So we can simply just have one such feature, let  $f_1 := f_{1a}$ . Next we can feature on how sleepy I am coming to the lecture.

$$f_2(x, y) = \begin{cases} 1 & \text{if } q_2(x) \geq 5 \wedge y = 1 \\ 0 & \text{otherwise} \end{cases}$$

A feature on whether the lecture was on zoom:

$$f_3(x, y) = \begin{cases} 1 & \text{if } q_3(x) = 0 \wedge y = 0 \\ 0 & \text{otherwise} \end{cases}$$

A feature on my background knowledge:

$$f_4(x, y) = \begin{cases} 1 & \text{if } y = 1 \\ 0 & \text{otherwise} \end{cases}$$

Let my feature vector be weight  $w \in \mathbb{R}^k$ . As of my experience, I almost get nothing out whenever its a zoom lecture. So I will give a high weight  $w_3$ . There should also be some weight to  $f_4$  as I still get something out. Here is another feature that I might consider. This feature is not *binary*. This is my cumulative knowledge increase as getting older:

$$f_5(x, y) = \begin{cases} q_5(x) & \text{if } y = 1 \\ 0 & \text{otherwise} \end{cases}$$

Let me remark on how one can easily enlarge the number of features to hundreds.

- my scale can be finer, going from 1-100.
- I can vary my constraints on features.

$$f(x, y) = \begin{cases} 1 & \text{if } \bigwedge_{i=1}^N \{q_i(x) = a_i\} \wedge y = 1 \\ 0 & \text{otherwise} \end{cases}$$

where  $\{a_i\}_{i=1}^N \in \mathbb{R}$ .

- Increase the value of  $N$ . In example, I have  $N = 5$ .

## Training data

Clearly, these are based on my experience. Perhaps for better model, incorporating other student's data might help.