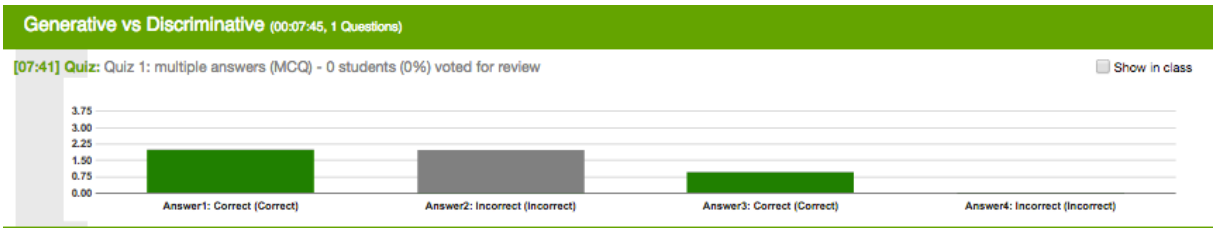
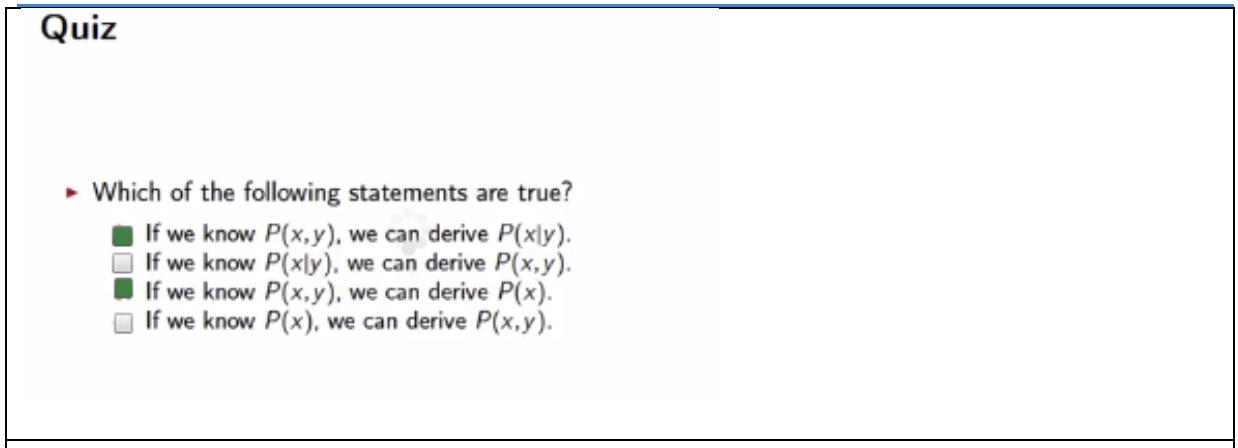
Monday 5 Dec 2016

Feedback: Quizzes

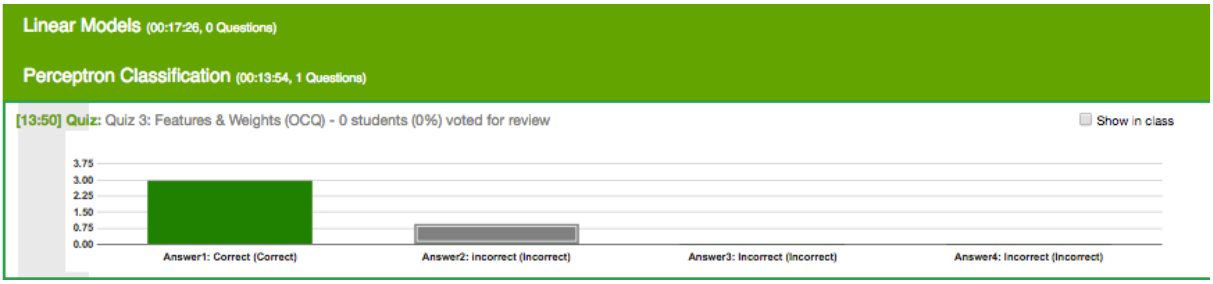
Quiz 1



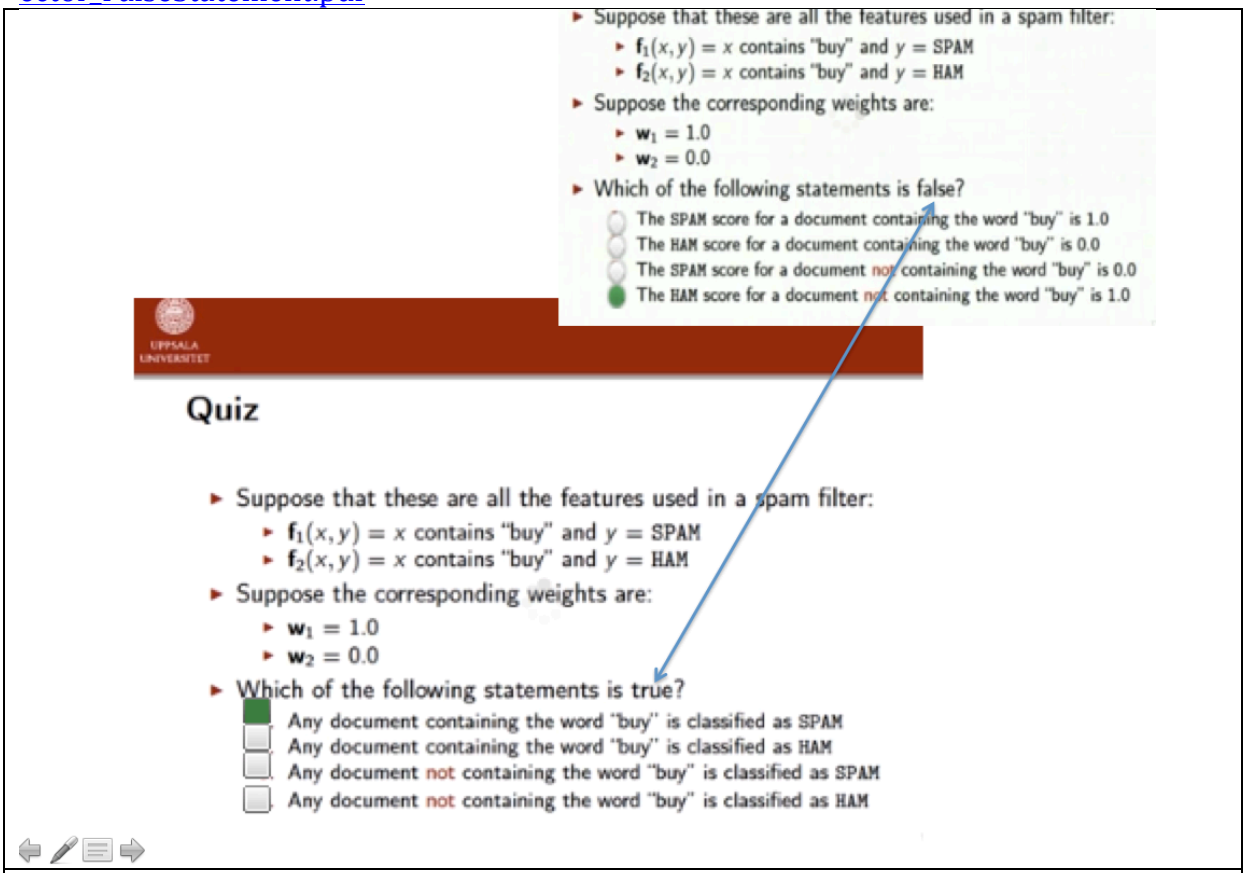


1: Correct: *P(x | y) = P(x, y) / Sum_y P(x, y)*
2: Incorrect: There is no way to derive P(x,y) from P(x|y).
3: Correct*: P(x) = Sum_y P(x, y)*
4: Incorrect: There is no way to derive P(x,y) from P(x).

Quiz 2



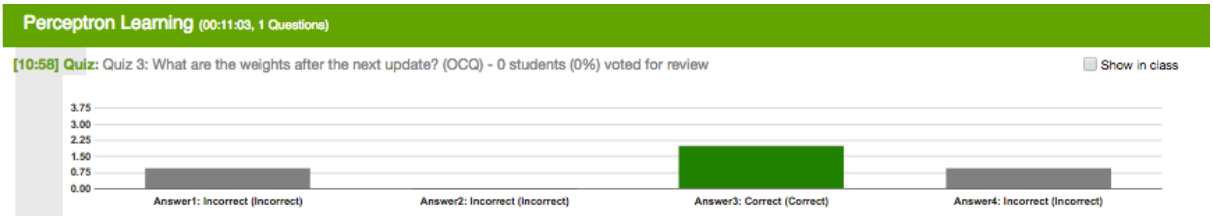
See calculations here:
http://stp.lingfil.uu.se/~santinim/ml/2016/Lect_07/07_Quiz2_LinearModels_FeatureVector_FalseStatement.pdf

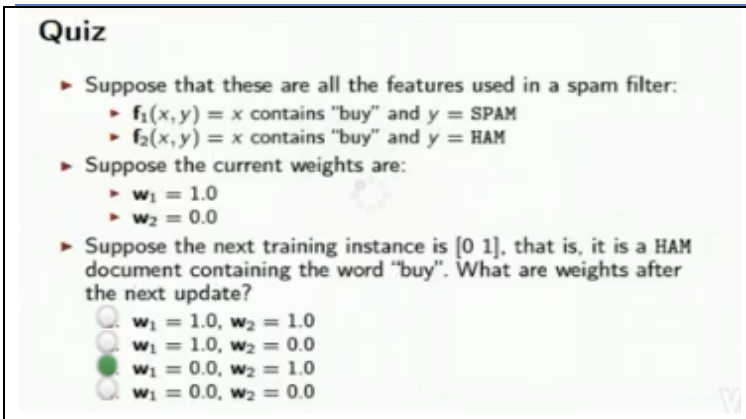


1. **Any document containing "buy" will get the score 1.0 for the SPAM class and 0.0 for the HAM class. (correct)**
2. Any document containing "buy" will get the score 1.0 for the SPAM class and 0.0 for the HAM class.
3. Any document NOT containing "buy" will get the score 0.0 for the SPAM class and 0.0 for the HAM class.

4. Any document NOT containing "buy" will get the score 0.0 for the SPAM class and 0.0 for the HAM class.

Quiz 3





1: Incorrect.
2: Incorrect.
3 Correct: We subtract 1 from the first weight and add 1 to the second weight.
4: Incorrect.

Listen again to this part of the videoclip:

Update Example

- Current weight vector:

  $\mathbf{w} = [1.0 \; -0.5 \; 0.0 \; -1.0 \; -2.0 \; 1.0 \; 0.5 \; 2.5]$

- Correct class:

  $\mathbf{f}(x, y) = [0 \; 0 \; 0 \; 0 \; 1 \; 1 \; 0 \; 1]$

- Predicted class:

  $\mathbf{f}(x, y') = [1 \; 1 \; 0 \; 1 \; 0 \; 0 \; 0 \; 0]$

- New weight vector

  $\mathbf{w} = [0.0 \; -1.5 \; 0.0 \; -2.0 \; -1.0 \; 2.0 \; 0.5 \; 3.5]$

- If after this replay, the weight calculations are still unclear, send me your personal calculations together with your reasoning, so I can see where you get stuck, and i will be able to help you further...

Remember:

Choose a **w** that minimizes error

$$\mathbf{w} = \arg\min_{\mathbf{w}} \sum_{t} 1 - \mathbb{1}[y_t = \arg\max_{y} \mathbf{w} \cdot \mathbf{f}(x_t, y)]$$

$$\mathbb{1}[p] = \begin{cases} 1 & p \text{ is true} \\ 0 & \text{otherwise} \end{cases}$$

This is a 0-1 loss function

The idea behind the perceptron is that we want to choose a weight vector **that minimize error on the training set.**
What does it mean? It means that we have **to find a value for w that minimizes this function**.
The function is simply this:

1. we sum over the entire training set and for each instance we compare the classification (this is the arg max given the weight vector w, this is the linear classification) of that point x sub t and this is y, and we compare that to the true classification in the training set, that is y sub t.
2. 1 is Logical function that means that whatever is see betw brackets is 1 or true and 0 otherwise. Whenever the classification is correct, then this thing is true, so the value of the function is 1 and then we subtract this to 1 and we get 0, so if the classification is correct we get 0 error. If it is incorrect we get one penalty point. And this is the function we want to minimize.
3. *it turns out that this a function that is very hard to minimize*

* * *