8. Prior, Posterior and Likelihood

Prior, Posterior and Likelihood

So here again, we will have multiple thetas

because we will have a big vocabulary, more than too many

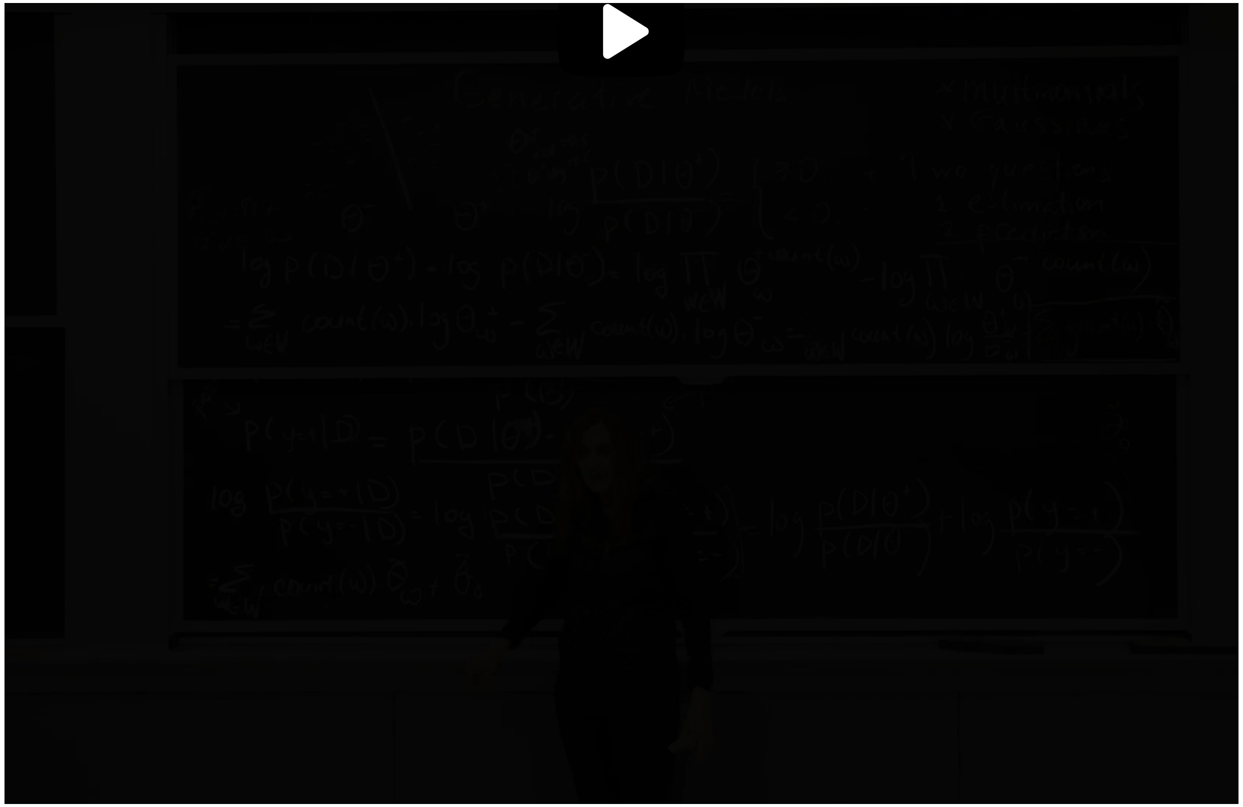
times, that can respond to the negative class,

and different kind of parameters of theta

for the positive class.

And the relation between them would determine

how this expression look like.



End of transcript. Skip to the start.

8:56 / 8:56 1.50x

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Prior, Posterior and Likelihood distributions

1/1 point (graded)

Consider a binary classification task with two labels '+' (positive) and '-' (negative).

Let y denote the classification label assigned to a document D by a multinomial generative model M with parameters θ^+ for the positive class and θ^- for the negative class.

Which of the following option(s) is/are true about the prior, posterior or likelihood distributions for this classifier? Choose the correct notations from the statements below:

- [x] $P(y = +|D)$ is the posterior distribution ✔
- [ ] $P(y = +|D)$ is the prior distribution
- [ ] $P(y = +)$ is the posterior distribution
- [x] $P(y = +)$ is the prior distribution ✔

✔

Solution:

Recall from the lecture that from bayesian rule we have,

$$P(y = +|D) = \frac{P(D|\theta^+) \times P(y = +)}{P(D)}$$

where $P(y=+|D)$ is the posterior distribution and $P(y=+)$ is the prior distribution while $P(D|\theta^{+})$ is the likelihood of document D given parameter θ^{+}

Submit You have used 1 of 1 attempt

Answers are displayed within the problem

A Numerical Example

2/2 points (graded)

Let's say that the prior for the positive class takes the following value:

$$P(y=+) = 0.3$$

Also, say that $P(D|\theta^{+}) = .3$ and $P(D|\theta^{-}) = .6$

From the above values of prior and likelihood, calculate the value of $P(D)$, the probability of generating document D. Enter the value below:

0.51

Answer: 0.51

From $P(D)$ also estimate the posterior probability $P(y=+|D)$. Enter your answer as a numerical expression or round it off to two decimal places.

9/51

Answer: 0.17647058823529413

Solution:

From bayes rule, we have that

$$P(D) = P(D|y=+)\,P(y=+) + P(D|y=-)\,P(y=-)$$

Also probability values must sum to 1 across all classes,

$$P(y=+) + P(y=-) = 1$$

Therefore,

$$P(D) = P(D|y=+) \times .3 + P(D|y=-) \times .7$$

$$P(D) = .3 \times .3 + .6 \times .7 = 0.51$$

From $P(D)$, we can calculate the posterior value $P(y=+|D)$ using bayes rule as follows:

$$P\left(y=+|D\right)=\frac{P\left(D|\theta^{+}\right)P\left(y=+\right)}{P\left(D\right)}$$

$$P\left(y=+|D\right)=\frac{.3\times .3}{0.51}=0.17647058823529 41$$

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 4 Unsupervised Learning (2 weeks) :Lecture 15. Generative Models / 8. Prior, Posterior and Likelihood

Add a Post

Show all posts ▼ by recent activity ▼

Fantastic connection and a suggestion
Congratulations to Prof. Barzilay and the staff for this brilliant set of lectures. Coming from the derivation of Bayes' theorem in 6.431x and seeing it a... 4

P(D)
Why isn't P(D) the sum of all probabilities P(D|+) + P(D|-)? what am I missing? 2