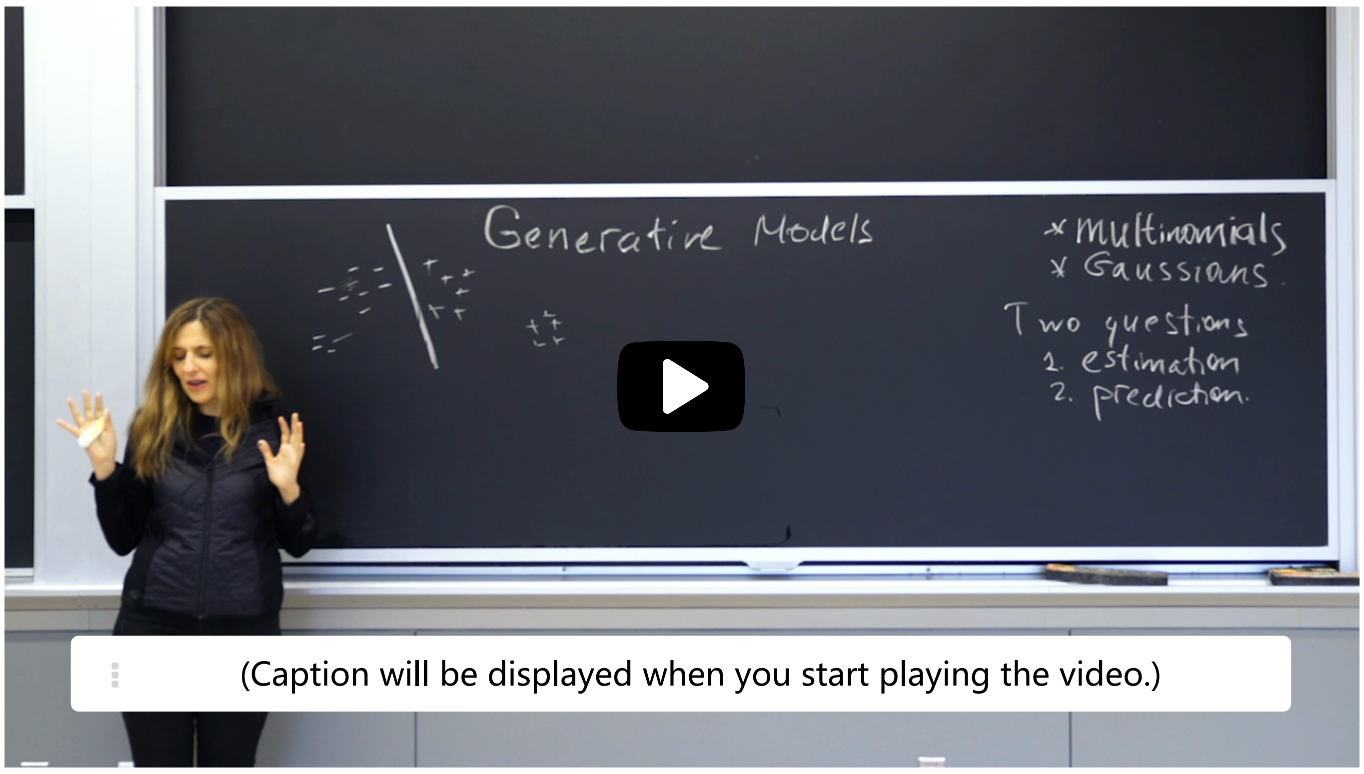
7. Prediction

Prediction

Start of transcript. Skip to the end.



0:00 / 0:00 1.0x

So now, we are ready to start looking
at the question of prediction.
So again, as the same way as in the case
of our discriminative supervised model,
we will have our points, let's say, just two classes, pluses
and minuses.
And using the estimation techniques,

Video
Download video file

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

Predictions of a generative multinomial model

1/1 point (graded)

Consider using a multinomial generative model M for the task of binary classification consisting of two classes which are denoted by + (positive class) and - (negative class).

Let the parameters of M that maximize the likelihood of training data for the positive class be denoted by θ^+ and for the negative class be denoted by θ^-.

Also, suppose that we classify a new document D to belong to the positive class iff

$$\log \frac{P(D|\theta^+)}{P(D|\theta^-)} \geq 0$$

where $P(D|\theta)$ stands for the probability that document D is generated using a multinomial distribution with parameters θ.

Which of the following option(s) is/are true about this generative classifier? Choose all that apply from the statements below:

- [x] A document is classified as positive iff $P(D|\theta^+) \geq P(D|\theta^-)$ ✔
- [ ] A document is classified as positive iff $P(D|\theta^+) < P(D|\theta^-)$

- [x] The generative classifier M can be shown to be equivalent to a linear classifier given by $\sum_{w \in W} count(w) \times \theta'_w \geq 0$ where $\theta' = \log \frac{\theta_w^+}{\theta_w^-}$ ✔

- [ ] The generative classifier M can be shown to be equivalent to a linear classifier given by $\sum_{w \in W} count(w) \times \theta'_w \geq 0$ where $\theta' = \log \frac{\theta_w^-}{\theta_w^+}$

✔

Solution:

Note that we classify a new document D to belong to the positive class iff $\log \frac{P(D|\theta^+)}{P(D|\theta^-)} \geq 0$ and to the negative class otherwise.

$$\log \frac{P(D|\theta^+)}{P(D|\theta^-)} \geq 0$$

is equivalent to

$$P(D|\theta^+) \geq P(D|\theta^-)$$

.

Recall from the lecture that,

$$\log \frac{P(D|\theta^+)}{P(D|\theta^-)}$$

$$= \log P(D|\theta^+) - \log P(D|\theta^-)$$

$$= \log \Pi_{w \in W} (\theta_w^+)^{count(w)} - \log \Pi_{w \in W} (\theta_w^-)^{count(w)}$$

$$= \sum_{w \in W} count(w) \log \theta_w^+ - \sum_{w \in W} count(w) \log \theta_w^-$$

$$= \sum_{w \in W} count(w) \log \frac{\theta_w^+}{\theta_w^-}$$

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 4 Unsupervised Learning (2 weeks) :Lecture 15. Generative Models / 7. Prediction