

Machine Learning Supervision 4

Third question is likely two questions on the first topic. Very unlikely to have two questions on HMM's. Could be a high level question "if you want to do machine learning how do you split your datasets" "what's the point of the development set" etc. Guesses. Fairly safe there will be something in Naïve Bayes, HMM's and perhaps more general machine learning methodology.

Sample

In machine learning a lot of things are "this would require empirical testing"

Note that you should answer what the question asks – they ask you about probabilities then think of probabilities.

Default is including states and then saying they're all zero.

"Showing both probabilities": include the probabilities. Can be shown by boxes with boxes and dotted lines showing the probability of meeting these states.

(c)

(i) The probability $a_{FL_2} = 0$ if no smoothing is used

(ii) In the training data L_2E will never occur and L_1E will never occur.

$$a_{L_1E}a_{L_2E} = 0$$

Note that all we can say about a_{FE} is that it's nonzero.

(iii) In the training data, we will not see $L_2L_2L_2$. with the first order model, we cannot model this since the next state only occurs on the current state. However, we could change the model to a second order HMM.

We can say $a_{L_2L_2} \leq \frac{1}{2}$.

(iv) Think about edge cases.

Naïve Bayes

Review binary features!!

(a)(i) Write the formula for naïve bayes and define the classes etc.

$$= \operatorname{argmax}_c P(x) \times \prod_i P(f_i|c)$$

$$c \in \{\text{igloo, house, tent, boat}\}$$

$$f_i = 1 \text{ if pixel } i \text{ is black} \\ 0 \text{ otherwise}$$

(ii) Read the question! How was the data labelled? How did you *get* the training data. Did you get people to annotate drawings or did you get people to draw houses and igloos etc.

(b) Explain better

(c) (i) Note they didn't ask you to *find* the features – just to say what the best would be.

(ii) Use the abstract feature classifier. If your confidence level is not great then backtrack to the raw pixel classifier.

If you just append features then they'll get swallowed by the 7500 other features.

General

Features having values:

If a feature is a continuous variable then how do you model it in naïve bayes.