Master M2 MSIAM and MoSIG

Fundamentals of Probabilistic Data Mining

Graded lab and homeworks

`http://chamilo.grenoble-inp.fr/courses/ENSIMAGWMM9AM017/`

2 Hidden Markov models

In this lab session we will be discussing hidden Markov models and their application to automatic speech recognition, and more precisely to phoneme recognition. We will be using the hidden Markov model toolkit (HTK: `http://htk.eng.cam.ac.uk/`). We will provide you the necessary tools to use HTK, the documentation and the data. The focus of this lab is to understand the behavior of the model(s) and algorithm(s). You have the following resources:

- The HTK documentation [`www.gipsa-lab.fr/~thomas.hueber/cours/htkbook.pdf`].
- The speech data pre-processed for use [provided later on].
- The compiled HTK routines [provided later on] (or the source code [provided later on]).
- The basis python script that we strongly suggest to use [provided later on].

2.1 Preliminary work

Do this before the class. Questions about this part will be answered only at the beginning of the practical session.

1. Describe the forward-backward algorithm of the EM for HMM. Recall the forward and backward recursions. What is this algorithm computing (provide a formula and an explanation)? What is the relationship between the state occupancy and the forward and backward recursions?

2. Describe the Viterbi algorithm. Recall its recursion. What is this algorithm computing (provide a formula and an explanation)?

3. What is the main difference between the forward and the Viterbi algorithms? Explain this in your own words.

4. In HTK, the EM for HMM (also called Baum-Welch) is implemented in the `HRest` routine. What are the main input and output variables of this routine?

5. As for any EM algorithm, the Baum-Welch algorithm needs to be initialised. In HTK, the initialisation routine is `HInit`. What are the main input and output variables of this routine? Describe this routine (both for the Gaussian and the GMM emission probability case). What is the rationale of this routine?

2.2 HMM with Gaussian emission probabilities

In this section we will be using HMMs with Gaussian emission probabilities.

2.2.1 Isolated training

1. Take a look to the provided script. How many hidden Markov models are we going to train?

2. Run the isolated training. Report the parameters of the model for vowel “a”. How many states are we using to model vowel “a”? And to model vowel “o”?

3. Launch the training. Identify and monitor the progress of the log-likelihood. Provide a plot for two different models.

4. Discuss the transition probabilities for “a” and “o”. Do you observe a pattern? Is the pattern repeated for other vowels?

5. Take a look to the model specifications in `models/proto`. Report the initialisation of all the parameters. Taking into account the initialisation of the weights, and the forward-backward and the re-estimation formulae of the transition probabilities: could you justify the pattern you observed for the transition probabilities? Provide a graphic representation of the state machine associated to this pattern.

6. Report the accuracy performance in the test set.

2.2.2 Connected training (embedded parameter re-estimation)

1. Launch the connected training using HERest routine. What is the difference between the connected and isolated training?

   Hint: let us consider the word “couleur”. Let us suppose that a few observations at the beginning of the sound “ou” are wrongly labeled as the end of sound “c”. What would happen in isolated training? And in connected training?

2. Report the parameters of the same models you reported in question 3). Do you observe any changes?

3. Report the accuracy performance in the test set. Discuss any differences with respect to the isolated training.

2.3 Mandatory additional questions

2.3.1 HMM with GMM emission probabilities

In this section we will be using HMMs with GMM emission probabilities.

1. What changes do you need to apply to `models/proto` so as to use a GMM as emission probability?

2. Run the isolated training for a low number of GMM components: is the log-likelihood still growing?

3. Report the accuracy performance in the test set. Discuss any differences with respect to the previous case (in isolated training).

4. Run the connected training with the same number of GMM components and report the accuracy. Discuss any differences with the previous cases.

5. Produce an accuracy curve as a function of the number of GMM components. Is the accuracy systematically growing with the number of components? Provide an explanation of the phenomenon you observe.

2.3.2 Full covariance matrices

1. What kind of covariance matrix have we used until now?

2. Does HTK allow the use of other types of covariance matrices? What happens if you train the models with full covariance matrices?

   Hint: follow the same path as in the previous sections. Start with Gaussian emission probabilities and isolated training, then connected training. Then switch to GMM emission probabilities and isolated training, the connected training. Finally increase the number of GMM components. Report the accuracy in all your experiments.

2.4 Optional additional questions

- We have seen that there are different choices in the model (# of GMM components, full/diagonal covariance matrix). Propose a protocol to automatically choose among these different models.

- Provide at least three examples of extensions of hidden Markov models, together with the reference to the published article.

- Provide the motivation (from the modeling perspective) of each of these variants.
- Sketch the differences in the E and/or M steps of the associated EM algorithm for each of these variants.