

Data Mining and Machine Learning

Assignment Project Exam Help

HMM Ad <https://eduassistpro.github.io/>

Add WeChat edu_assist_pro

Peter Jančovič



Objectives

- So far we talked about Maximum Likelihood training for HMMs (the E-M algorithm)
 - Viterbi-style training
 - Baum-Wel
- In this session, we talk about adaptation:
 - Maximum A-Posteriori (MAP) estimation
 - Maximum Likelihood Linear Regression (MLLR)



Adaptation

- A modern large-vocabulary continuous speech recognition system has many thousands of parameters
- Many hours (e.g. 200+ h) train the system
- Speech data comes from
- Hence recogniser is ‘speaker independent’
- But performance for an individual would be better if the system were speaker dependent

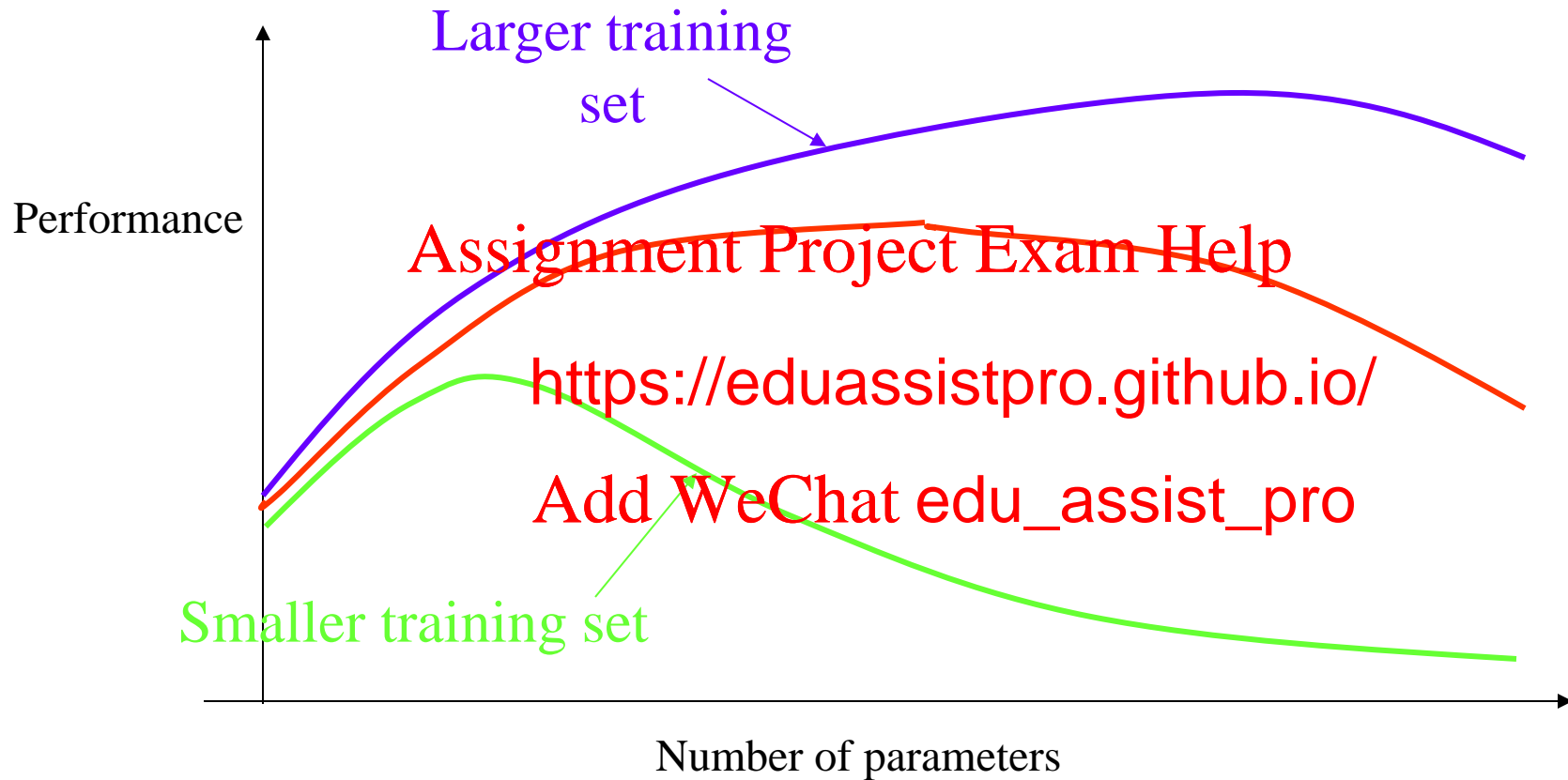


Adaptation

- For a single speaker, only a small amount of training data is available
- Viterbi reestimation or Baum-Welch reestimation will not work <https://eduassistpro.github.io/>
- Adaptation: **Add WeChat edu_assist_pro**
 - the problem of robustly adapting a large number of model parameters using a small amount of training data



‘Parameters vs training data’



Adaptation

- Two common approaches to adaptation (with small amounts of training data)
 - Bayesian adaptation (also known as MAP adaptation)
(MAP = Maximum a Posteriori)
 - Transform-based adaptation (also known as MLLR)
(MLLR = Maximum Likelihood Linear Regression))

Assignment Project Exam Help

<https://eduassistpro.github.io/>

Add WeChat: edu_assist_pro



Bayesian (MAP) adaptation

- MAP estimation maximises the posterior probability of M given the data y , i.e., $P(M | y)$
- From Bayes' Theorem:

Assignment Project Exam Help

<https://eduassistpro.github.io/>

Add WeChat edu_assist_pro

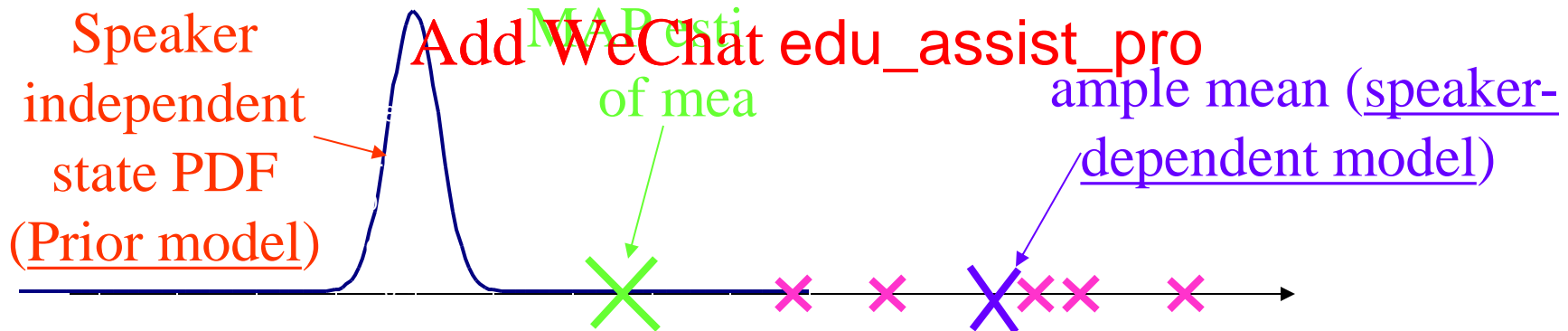
- $P(M)$ is the prior probability of M
- $p(y / M)$ is the likelihood of the adaptation data on M



Bayesian (MAP) adaptation

- Uses well-trained, ‘speaker-independent’ HMM as a prior $P(M)$ for the estimate of the parameters of the speaker dependent HMM
- E.G:

<https://eduassistpro.github.io/>



Bayesian (MAP) adaptation

$$\hat{M} = \lambda M_{prior} + (1 - \lambda) M_y, 0 \leq \lambda \leq 1$$

MAP model Prior model Speaker-dependent'

<https://eduassistpro.github.io/>

- Intuitively, if the adaptation data is large, then the MAP adapted model will be biased towards y , so λ will be small
- Conversely, if there is very little adaptation data, the MAP model will be biased towards the prior, so λ will be big



Transform-based adaptation (MLLR)

- Maximum Likelihood Linear Regression (MLLR) is another method for adapting the mean vectors of a set of HMMs
- Estimate a linear transform to transform speaker-independent into speaker-dependent parameters
- Suppose that <https://eduassistpro.github.io/> with Gaussian Mixture state output
- Suppose A is linear transform d -dimensional space of acoustic vectors and that b is an acoustic vector
- Let $M_{SD} = T(M_{SI})$ be the HMM derived from M_{SI} by replacing each Gaussian mean vector μ with $A\mu + b$



MLLR adaptation

- Given data y from a new speaker, the aim of MLLR is to find A and b such that $P(y/T(M_{SI}))$ is maximised
- ... hence Maximum Likelihood LR
- Need to estimate the $D \times D$ parameters of A
- Each acoustic transformation of the original, so a linear transform of the 40 = 1600 parameters
- This is much less than the 10s thousands of parameters needed to train the model
- Same transformation A can be used for all models and states.
- Alternatively, if there is enough data from the new speaker, a separate transformation can be estimated for each model, state, or set of states



Transform-based adaptation

Speaker-independent
parameters

Speaker-dependent
data points

Assignment Project Exam Help

<https://eduassistpro.github.io/>

Add WeChat edu_assist_pro

Adapted parameters

‘best fit’
transform



Summary

- Bayesian (MAP) adaptation
 - J-L Gauvain and C-H Lee, “Bayesian learning for Hidden Markov Models with Gaussian mixture state observation densities”, *Speech Communication* 11, pp 205-213, 1992
- Transform-bas <https://eduassistpro.github.io/>
 - C J Leggetter and P C Woodland, “Maximum likelihood linear regression for speaker adaptation of density HMMs”, *Computer Speech and Language*, 1995

