Neural networks
Training CRFs - discriminative vs. generative learning
**Topics:** discriminative learning, generative learning

- In discriminative learning, we optimize the conditional likelihood: \(- \log p(y|X)\)
  - CRFs are discriminative
- In generative learning, we optimize the joint log-likelihood:
  \[- \log p(y, X) = - \log (p(y|X)p(X)) = - \log p(y|X) - \log p(X)\]
  - HMMs are usually trained generatively
  - \(- \log p(X)\) is similar to a regularizer
Topics: generative learning, discriminative learning

• It can be shown that:
  ‣ if model is well-specified (i.e. is the true model) generative learning is better
Topics: generative learning, discriminative learning

• It can be shown that:
  ‣ if model is not well-specified (i.e. most of the time), it depends:
  ‣ See these papers for more details: