Applied Machine Learning

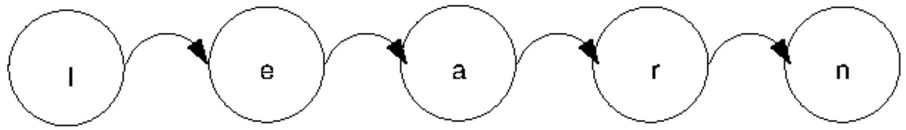
Hidden Markov Models

Hidden Markov Models

- Markov Chains and Hidden Markov Models (HMMs)
- Emission Distribution in HMMs
- Inference in HMMs

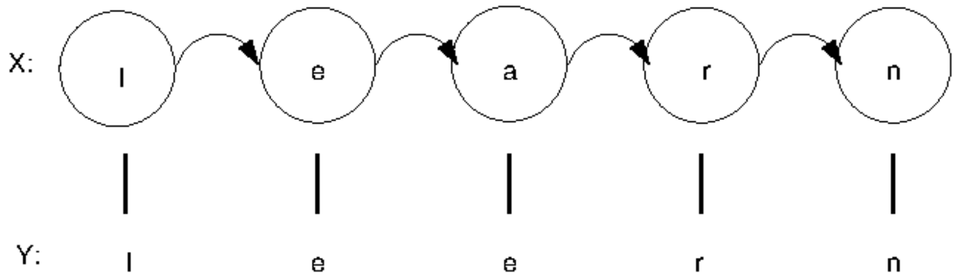
Hidden Markov Models

- Text input from a keyboard or touchscreen
  - Typos
- In general
  - Noise
- Hidden Markov Models (HMMs)
  - Set of states with transition probabilities
  - Initial distribution of states
  - Emission distribution



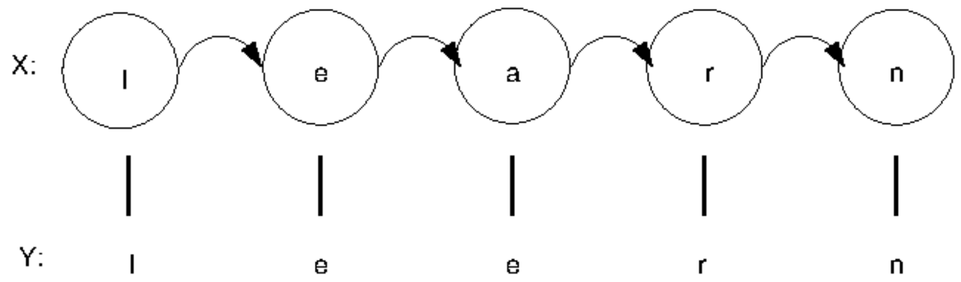
Hidden Markov Models

- Finite State, Time Homogeneous Markov Chain:
  - S states
  - Hidden state at step n: X_n
  - Transition Probability Matrix P:
    - $p_{i,j} = P(X_n = j \,|\, X_{n-1} = i)$
  - Initial distribution of states $\pi = P(X_1 = i)$
- Emission Distribution
  - at hidden state X_n, we observe Y_n
    - discrete random variable, O possible states
  - $q_i(Y_n) = P(Y_n \,|\, X_n = i)$: matrix $\mathbb{Q}$
- HMM: $(X, Y, P, \mathbb{Q}, \pi)$



Hidden Markov Models

- Speech transcription
  - M-grams for phonemes in spoken language
  - States: M-grams of phonemes
  - Transitions P: extract frequencies from datasets
  - Emission distribution $\mathcal{Q}$: characterize signals that result from spoken words
  - Initial distribution π



HMMs: Inference

- Inference
  - Estimate sequence of hidden states X_i for known HMM for sequence of observations Y_i
- Maximum a Posteriori Inference or MAP Inference
  - sequence of hidden states $X_1, X_2, \ldots, X_N$
  - maximize posterior
    - $P(X_1, X_2, \ldots, X_N \mid Y_1, Y_2, \ldots, Y_N, P, \mathcal{Q}, \pi)$
    - $= \frac{P(X_1, X_2, \ldots, X_N, Y_1, Y_2, \ldots, Y_N \mid P, \mathcal{Q}, \pi)}{P(Y_1, Y_2, \ldots, Y_N)}$
- Same as minimizing
  - $-\log \left(\frac{P(X_1, X_2, \ldots, X_N, Y_1, Y_2, \ldots, Y_N \mid P, \mathcal{Q}, \pi)}{P(Y_1, Y_2, \ldots, Y_N)} \right)$
  - $= -\log P(X_1, X_2, \ldots, X_N, Y_1, Y_2, \ldots, Y_N \mid P, \mathcal{Q}, \pi) + \log P(Y_1, Y_2, \ldots, Y_N)$
- Cost Function to Minimize:

$$-\log P(X_1, X_2, \ldots, X_N, Y_1, Y_2, \ldots, Y_N \mid P, \mathcal{Q}, \pi)$$

$$\begin{aligned} = -(& [\log P(X_1) + \log P(Y_1 \mid X_1)] + \\ & [\log P(X_2 \mid X_1) + \log P(Y_2 \mid X_2)] + \\ & \ldots + \\ & [\log P(X_N \mid X_{N-1}) + \log P(Y_N \mid X_N)] \\ &) \end{aligned}$$

Hidden Markov Models

- Markov Chains and Hidden Markov Models (HMMs)
- Emission Distribution in HMMs
- Inference in HMMs

Applied Machine Learning

Graphical Models I