Generative Models

Lyle Ungar

Generative models + EM estimation
GMM
Naïve Bayes
LDA
HMM

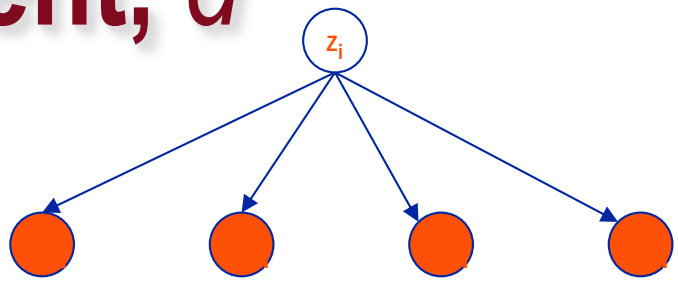
GMM - generation

- **To generate each point:**
  - Select a cluster *k* with probability π
  - Generate a point from the Gaussian distribution $N(\mu_k, \Sigma_k)$

GMM - estimation

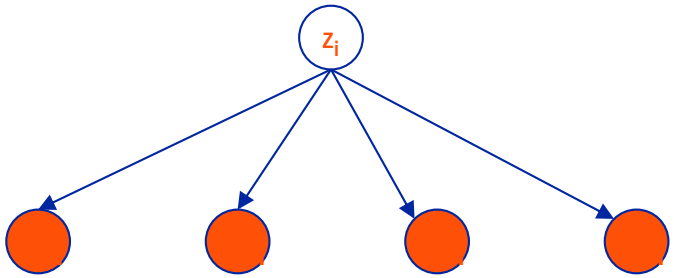
- E: estimate cluster assignments
- M: estimate μ, Σ

Naïve Bayes: To generate a document, *d*



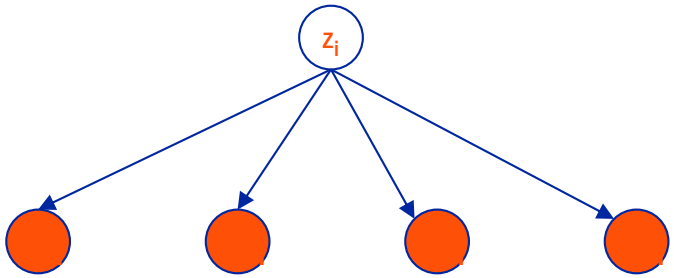
- Choose a topic z_d with $p(topic_i) = \theta$
- Choose N words w_n by drawing each one independently from a multinomial conditioned on z_d with $p(w_n=word_j|topic_i=z) = \beta_z$

Naïve Bayes: Estimation



- **Usually, the topic is observed, so just count**
  - $p(topic_i) = \theta$
  - $p(w_n=word_j|topic_i=z) = \beta_z$
- **Or use MAP**

Naïve Bayes: Estimation



- **If the topic is not observed, use MLE**
  - *E:* $\theta = p(topic_i)$
  - *M:* $= \beta_z = p(w_n = word_j | topic_i = z)$

LDA: To generate a document, *d*

- Choose a *multinomial distribution* θ_d ~ *Dirichlet*(α) over topics for that document
- For each of the *N* words w_n in the document
  - Choose a topic z_n ~ Multinomial(θ_d)
  - Choose a word w_n ~ Multinomial(β_z) from that topic z_n
    - $\beta_z = p(w=w_n|topic=z_n)$

LDA: Estimation

- **E-step:**
  - Compute $p(\theta, \mathbf{z} | \mathbf{w}, \alpha, \beta)$, the posterior of the hidden variables $(\theta, \mathbf{z})$ given each document $\mathbf{w}$, and parameters α and β.
    - θ_{dk} = *p(topic k in document* d*)*
    - z_{dn} = *topic of word n in document d*
- **M-step**
  - Estimate parameters α and β given the current hidden variable distribution estimates
    - β_{kv} = *p(w=v | topic=k)*
    - α = prior for Dirichlet (can also be a hyperparameter)

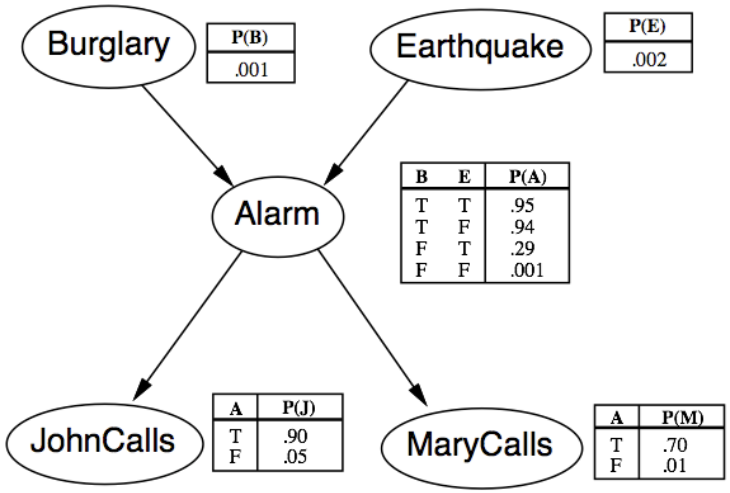
HMM - generation

- **To generate each observation:**
  - Select an initial hidden state s_0 with probability π
  - Repeat
    - Generate an emission with probability $p(x_i|s_i)$
    - Transition to the next state with probability $p(s_{i+1}|s_i)$

HMM - estimation

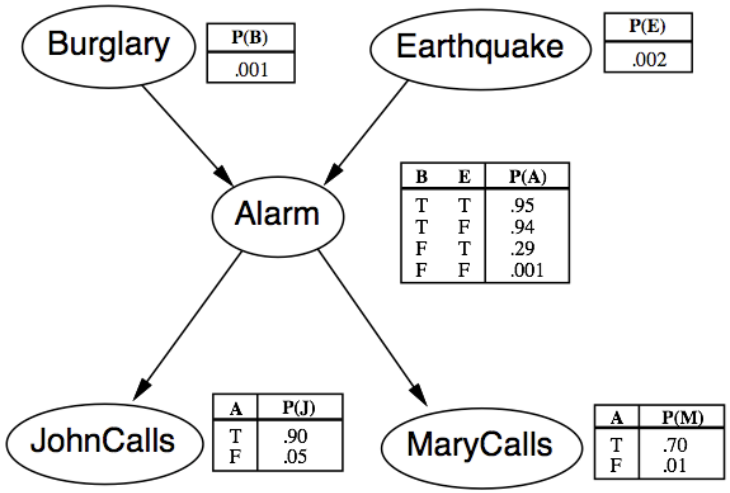
- E: estimate hidden state
- M: estimate emission and transition probabilities

Bayes Net Generation



- *Select values for the top level features with probability…*
- *Repeat*
  - *Select values for the next layer down, given the probability tables*

Bayes Net Generation



- ***Write the network structure***
- ***Estimate the probabilities by counting***