Applied Machine Learning

Markov Chains

Markov Chains

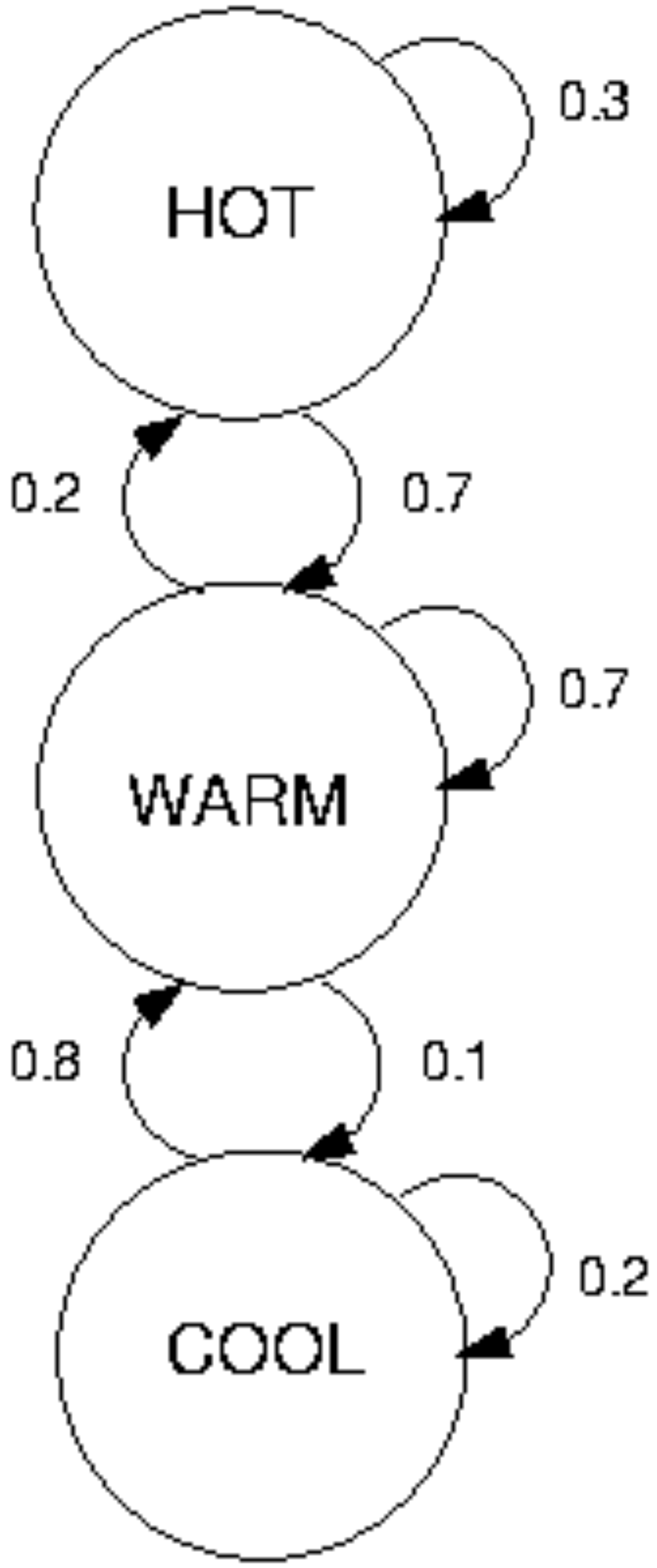
- Structure of Markov Chains
- Production of Markov Chains
- Components of Markov Models

Markov Chains

- Sequence:
    - [HOT, WARM, HOT, WARM, WARM, WARM]
- Sequence of random variables X_n
    - finite number of states X
- Transition probabilities:
    - $P(x_n = j | X_{n-1} = i)$
- Markov property
    - $P(X_n = j | X_{n-1}, X_{n-2}, \ldots X_0) = P(X_n = j | X_{n-1})$
- Discrete Time, Finite State, Time Homogeneous Markov Chain

Markov Chains as Biased Random Walks

- Finite Directed Graph
  - Edges: P(x_n =j | X_{n-1}=i) $\sum_j P(x_n = j \mid X_{n-1} = i) = 1$
- Markov Chains:
  - Biased Random Walk
    - [HOT, WARM, HOT, WARM, WARM, WARM]
    - [COLD, COLD, WARM, COLD, WARM]
  - There may be a start state or an initial distribution
  - Absorbing state: $P(X_n = j \mid X_{n=1} = j) = 1$
  - Recurrent state
    - may happen repeatedly in the same sequence
- Model processes that generate these Markov Chains
- Probability of arriving at some state



Markov Chains: States and Transition Probabilities

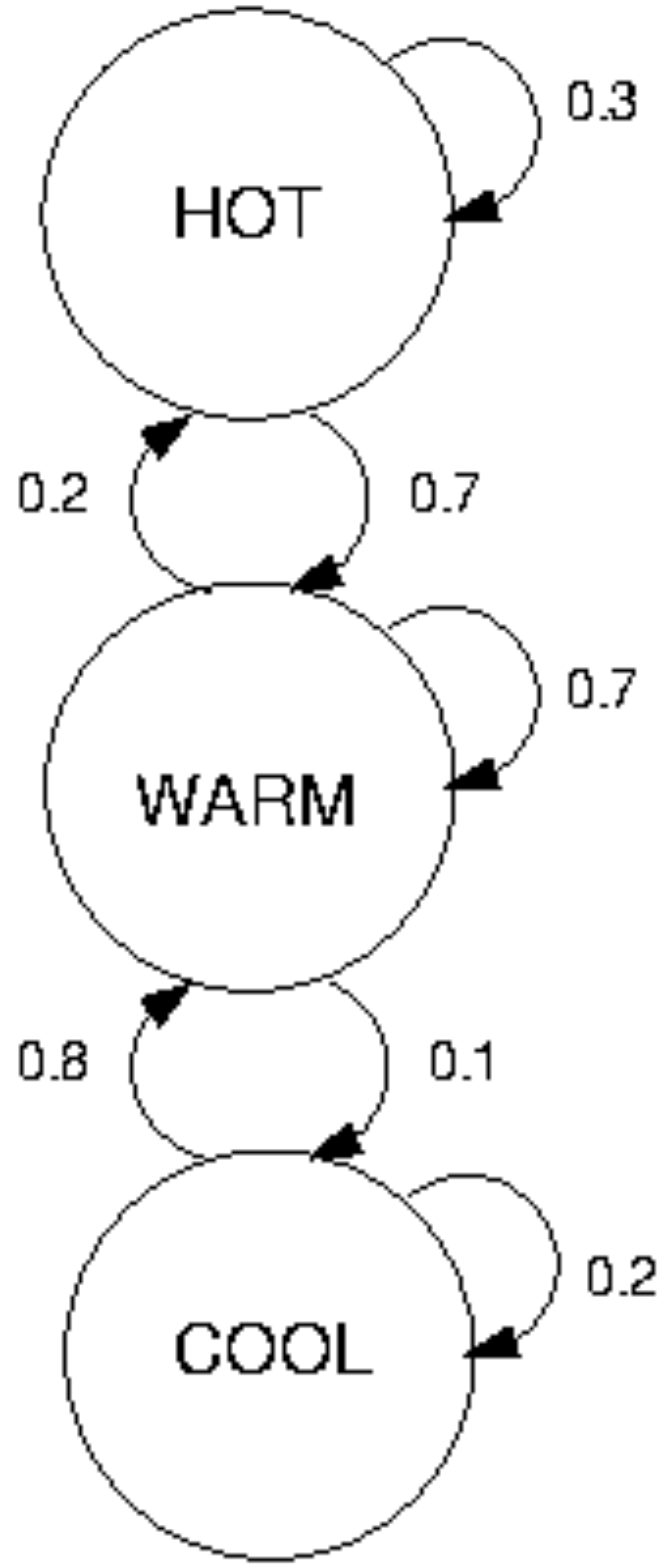
- State encoding:
  [hot $= 0$, warm $= 1$, cool $= 2$]

- Transition Probability Matrix P

  - $p_{i,j} = P(X_n = j \mid X_{n-1} = i)$

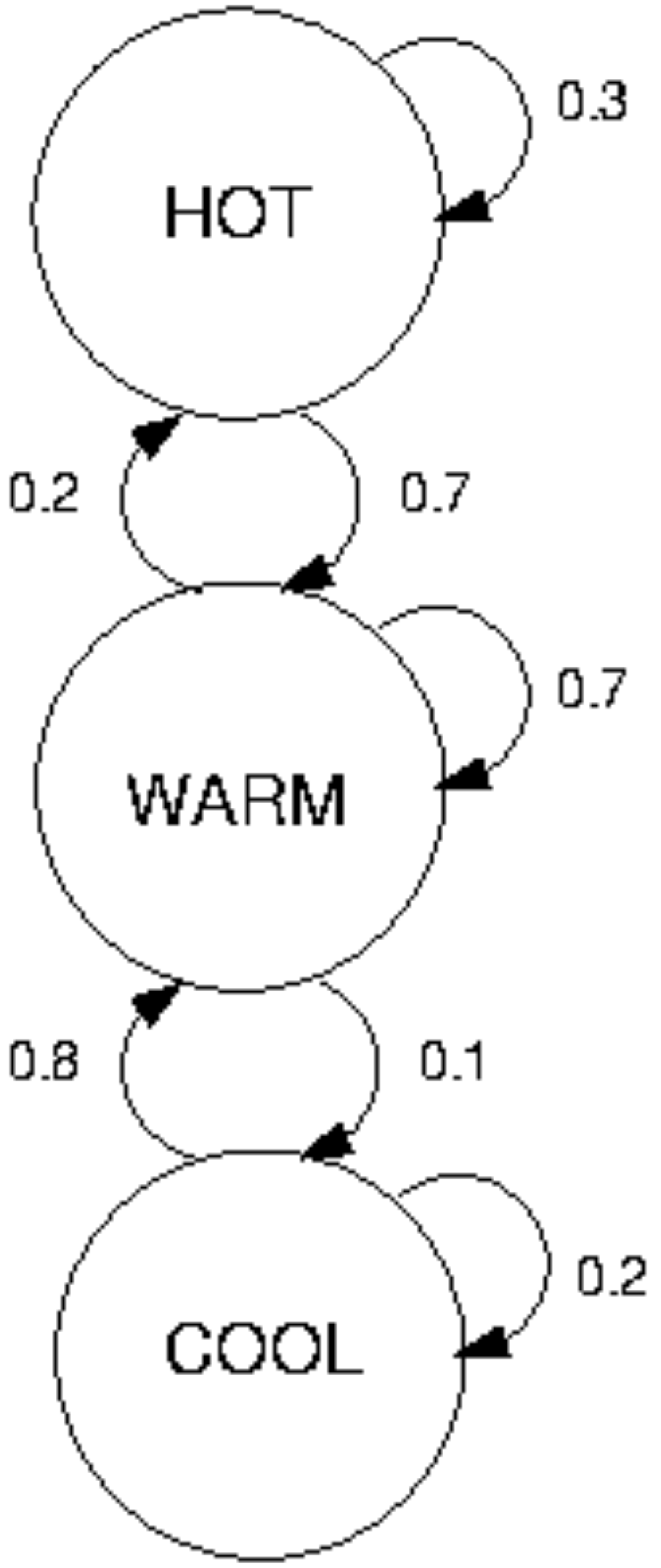
  - $P = \begin{bmatrix} 0.3 & 0.7 & 0 \\ 0.2 & 0.7 & 0.1 \\ 0 & 0.8 & 0.2 \end{bmatrix}$

  - $\sum_j P(x_n = j \mid X_{n-1} = i) = \sum_j p_{i,j} = 1$



Simulating Probability Distributions

- Initial probability distribution $P(X_0 = i)$
  - $\pi_{[1\times k]}$
  - $\pi = [0.0, 0.20, 0.80]$
- Probability distribution at time 1:
  - $$\begin{aligned} P(X_1 = j) &= \sum_i P(X_1 = j, X_0 = i) \\ &= \sum_i P(X_1 = j \mid X_0 = i) P(X_0 = 1) \\ &= \sum_i p_{i,j} \pi_i \end{aligned}$$
  - As matrix multiplication: $\pi_1 = \pi P$
  - $\pi_1 = [0.04, 0.78, 0.18]$



Markov Chains

- Structure of Markov Chains
- Production of Markov Chains
- Components of Markov Models

Applied Machine Learning

Markov Chains