



BITS Pilani

Pilani | Dubai | Goa | Hyderabad

Computational and Artificial Intelligence

Tutorial 6

Hidden Markov Models

Definition

- Hidden Markov Models (HMMs) are a class of probabilistic graphical model.
- They are used to predict a sequence of unknown (hidden) variables from a set of observed variables.
- An HMM can be viewed as a Bayes Net unrolled through time with observations made at a sequence of time steps being used to predict the best sequence of hidden states.
- A simple example of an HMM is predicting the weather (hidden variable) based on the type of clothes that someone wears (observed).

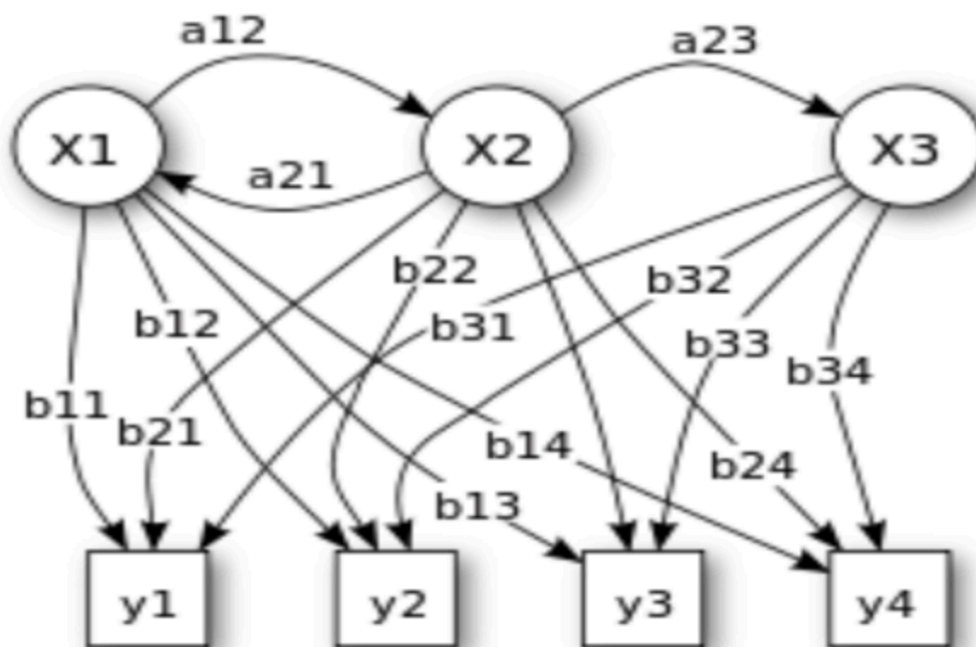


Figure 1. Probabilistic parameters of a hidden Markov model (example)

X — states

y — possible observations

a — state transition probabilities

b — output probabilities

The below diagram from Wikipedia shows an HMM and its transitions.

The scenario is a room that contains urns **X_1 , X_2 and X_3** , each of which contains a known mix of balls, each ball labeled **y_1 , y_2 , y_3 and y_4** .

A sequence of four balls is randomly drawn. In this particular case, **the user observes a sequence of balls y_1, y_2, y_3 and y_4 and is attempting to discern the hidden state which is the right sequence of three urns that these four balls were pulled from.**

Why Hidden Markov Model?

- The reason it is called a Hidden Markov Model is because we are constructing an inference model based on the assumptions of a **Markov process**.
- The Markov process assumption is simply that the “**future is independent of the past given the present**”. In other words, assuming we know our present state, we do not need any other historical information to predict the future state.
- Generally, the term “states” are used to refer to the hidden states and “observations” are used to refer to the observed states.

Important Terminologies

Transition data — the probability of transitioning to a new state conditioned on a present state.

Emission data — the probability of transitioning to an observed state conditioned on a hidden state.

Initial state information — the initial probability of transitioning to a hidden state. This can also be looked at as the prior probability.

HMM and POS Tagging

Scoring a set of sequences given some observations

Assuming that we need to determine the parts of speech tags (hidden state) given some sentence (the observed values), we will need to first score every possible sequence of hidden states and then pick the best sequence to determine the parts of speech for this sentence.

We will score this using the below steps

1. Generate the initial, transition and emission probability distribution from the sample data.
2. Generate a list of all unknown sequence
3. Score all unknown sequences and select the best sequence