

Date...17-Sept-2025...

Hidden Markov Model (HMM): HMM is a graphical probabilistic model used to predict a sequence of unknown values from a set of observed variables.

For HMM, we consider:

- (i) A set of states : $\{S_1, S_2, \dots, S_N\}$
- (ii) Process moves from one state to another generating a seq. of states. s.e.:

$\{S_1, S_2, \dots, S_k, \dots\}$

- (iii) Markov Chain Property : $P(S_k | S_1, S_2, \dots, S_{k-1})$.

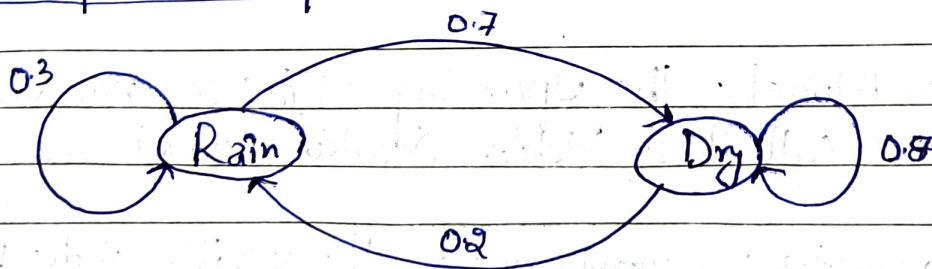
ie. The probability of current state depends only on the previous states (all previous).

- (iv) To define Markov Model, we consider some probabilities:

(a) Transition Probabilities : $a_{ij} = P(S_i | S_j)$

(b) Initial Probabilities : $\pi_i = P(S_i)$
↳ of each state.

Graphical Example:



Two States : $\{ \text{Rain, Dry} \}$

Transition Probabilities:

- $P(\text{Rain} | \text{Rain}) = 0.3$
- $P(\text{Rain} | \text{Dry}) = 0.2$
- $P(\text{Dry} | \text{Rain}) = 0.7$
- $P(\text{Dry} | \text{Dry}) = 0.8$

Initial Probabilities: (Say)

$$P(\text{Rain}) = 0.4$$

$$P(\text{Dry}) = 0.6$$

Q: How to calculate Sequence Probabilities?

Ans: for this we use Markov Chain Property.

$$P(S_1, S_2, \dots, S_k) = P(S_k | S_1, S_2, \dots, S_{k-1}) \cdot P(S_1, S_2, \dots, S_{k-1}).$$

$$= P(S_k | S_{k-1}) \cdot P(S_1, S_2, \dots, S_{k-1}).$$

$$= P(S_k | S_{k-1}) \cdot P(S_{k-1} | S_{k-2}) \cdot P(S_1, S_2, \dots, S_{k-2}).$$

$$\dots$$

$$= P(S_k | S_{k-1}) \cdot P(S_{k-1} | S_{k-2}) \cdot \dots \cdot P(S_2 | S_1) \cdot P(S_1).$$

Q: Calculate prob. of a sequence of states: {Dry, Dry, Rain, Rain}

Ans:

$$P(\text{Dry, Dry, Rain, Rain}) = P(\text{Rain} | \text{Rain}) * P(\text{Rain} | \text{Dry}) * P(\text{Dry} | \text{Dry}) * P(\text{Dry})$$

$$= 0.3 * 0.2 * 0.7 * 0.6$$

$$= 0.0288$$

Hidden Markov Model: The upcoming states are not visible and that's why we call it a hidden Markov Model.

- * Each state generates randomly one of the visible states based on M observations e.g. $\{v_1, v_2, \dots, v_M\}$.
- * The process moves from one state to other generating a sequence of states.

Date.....

Matrix of Transition Probabilities:

$$A = [a_{ij}], \text{ where } a_{ij} = P(S_i | S_j)$$

Matrix of Observation Probabilities:

$$B = [b_i(V_m)], \text{ where } b_i(V_m) = P(V_m | S_i)$$

Vector of Initial Probabilities:

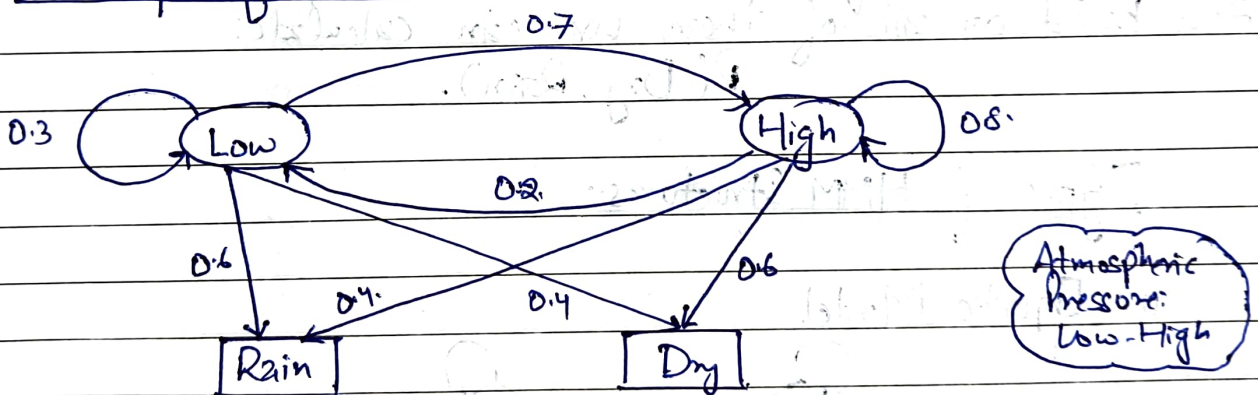
$$\pi = \pi_i, \text{ where } \pi_i = P(S_i)$$

Then,

$$HMM = (A, B, \pi)$$

A particular HMM can be represented using these properties.

Example of HMM



Transition Probabilities

$$P(\text{low} | \text{high}) = 0.3$$

$$P(\text{high} | \text{low}) = 0.7$$

$$P(\text{high} | \text{high}) = 0.8$$

$$P(\text{low} | \text{low}) = 0.2$$

Observation Probabilities

$$P(\text{Rain} | \text{low}) = 0.6$$

$$P(\text{Dry} | \text{low}) = 0.4$$

$$P(\text{Rain} | \text{high}) = 0.4$$

$$P(\text{Dry} | \text{high}) = 0.6$$

Initial Probabilities

$$P(\text{low}) = 0.4$$

$$P(\text{high}) = 0.6$$

Spiral

Teacher's Sign

Date.....

Calculation of Sequence Probability:

Q.1: ~~$P(\{Dry, Rain\})$~~ ~~$P(Dry, Rain)$~~ ~~$P(Rain)$~~

$$P(\{Dry, Rain\}) = P(\{Dry, Rain\} | \{low, low\}) \cdot P(\{Dry, Rain\} | \{low, high\}) \cdot P(\{Dry, Rain\} | \{high, low\}) \cdot P(\{Dry, Rain\} | \{high, high\})$$

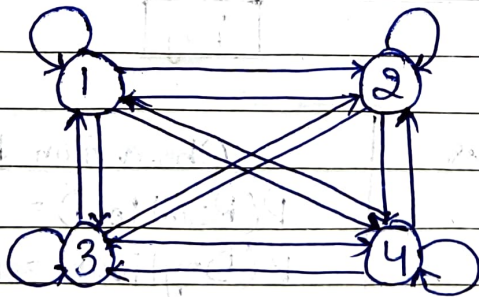
Now,

$$\begin{aligned} P(\{Dry, Rain\} | \{low, low\}) &= P(Dry | Rain) \cdot P(Rain | low) \cdot P(low | low) \cdot P(low) \\ &= 0.4 \times 0.6 \times 0.3 \times 0.4 \\ &= 0.0192 \end{aligned}$$

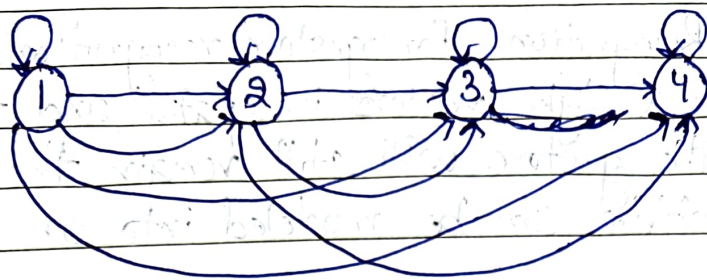
Similarly, we can find all other conditional probabilities as well. And, based on all of them, we can calculate:
 $P(Dry, Rain)$.

Types of HMM Structures:

(i) Ergodic Model



In ergodic model, every a_{ij} (transition probability) is +ve and every transition is possible.

(ii) Bakis (Left to Right) Model

* In this model, the $a_{ij} = 0 \forall j < i$ (i.e. backward transitions are not possible).

* In speech recognition, this is very popular model.

Main Issues in HMM (How can HMM work).

(i) Evaluation Problem: Given the model: $M = (A, B, \pi)$ and observation sequence $O = o_1, o_2, \dots, o_K$. Calculate the probability that the model M has generated the sequence O . (Forward-Backward Algorithm)

(ii) Decoding Problem: Given $M = (A, B, \pi)$ & the observation sequence $O = o_1, o_2, \dots, o_K$. Calculate the most likely sequence of hidden states S_i that produced this observation sequence O . (Viterbi Algorithm).

(iii) Learning Problem: Given some training observation sequences: $O = o_1, o_2, \dots, o_K$ and general structure of HMM (no. of hidden and visible states), determine HMM parameters. i.e. $M = (A, B, \pi)$ that best fit the training data.

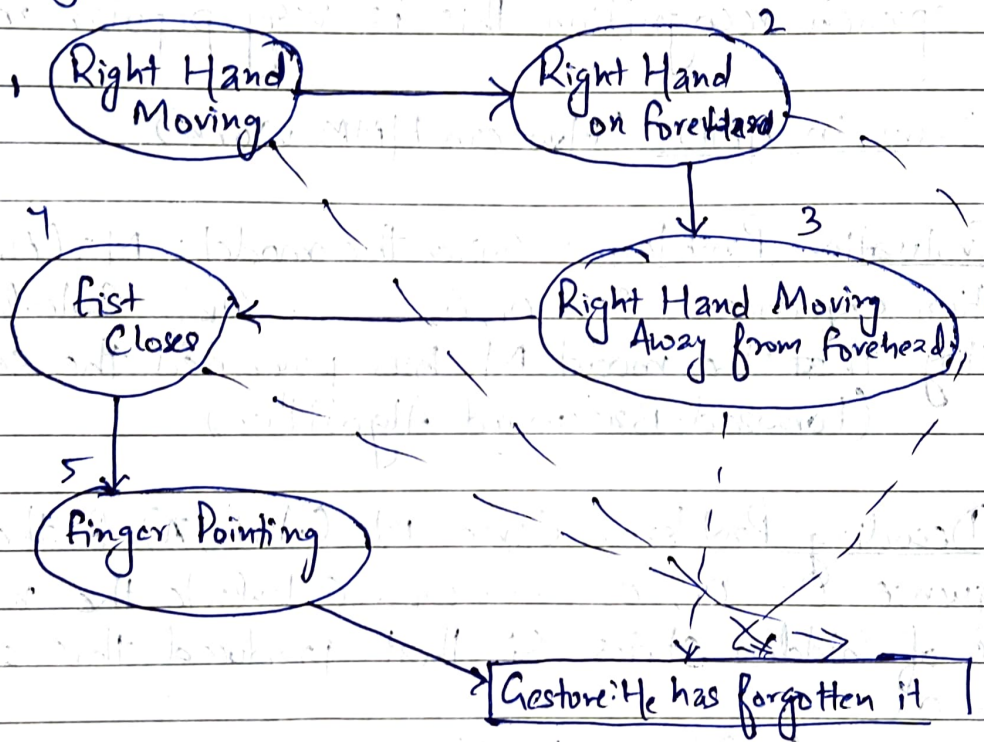
(Baum-Welch Algorithm).

Real Time Example for HMM Use

i) Gesture Recognition: For gesture recognition, different person can perform a slight variations in spatial and temporal form. However, the gesture will still remain the same. As a result, such problem can be modeled into an HMM.

e.g.

for Gesture: He has forgotten it, we can define the following states:



Based on this HMM, irrespective of whether there is a delay in the gesture in terms of spatial or temporal, if transition happens from 1 to 5 states, the gesture is predicted correctly. Thus, HMM models the temporal dependency in the learning.

* The current example is a learning problem of HMM.