

# HMMs parameter matrix calculation

How can one compute the emission and transition matrices for a Hidden Markov Model (HMM) based on a 3-character sequence? Additionally, could you provide the relevant formulas for these calculations?







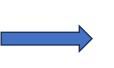
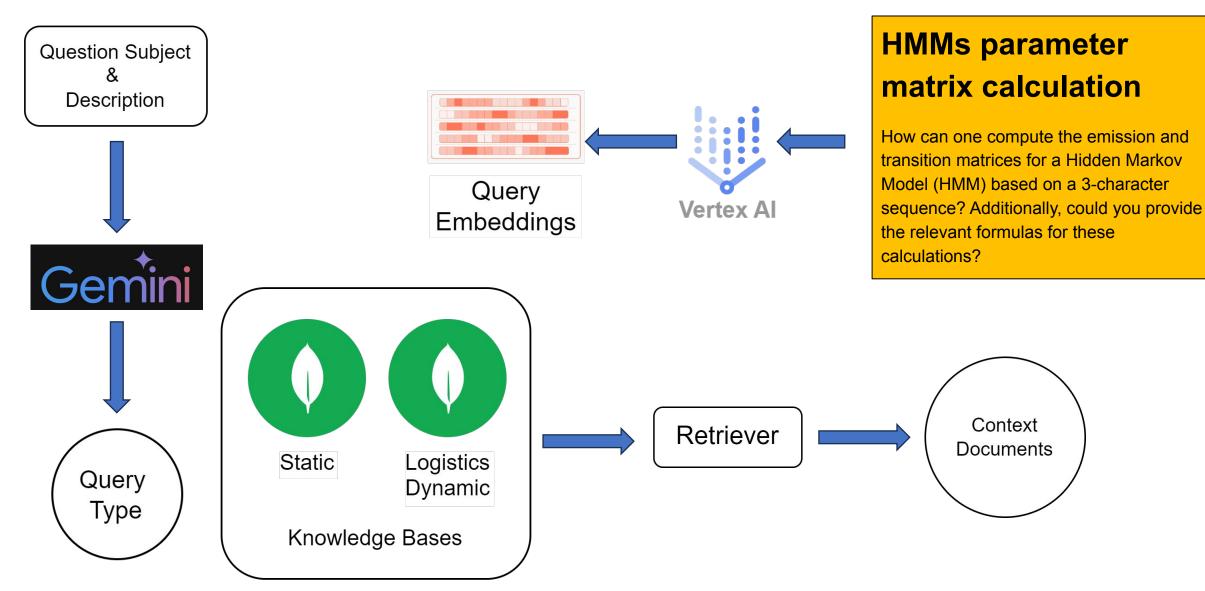


Image summary



**Images** 



## Semantic Similarity



## Homogenous HMM

· We include an additional assumption

$$P(S_j = s_j | S_{j-1} = s_{j-1}) = t(s_j | s_{j-1})$$

$$P(X_j = x_j | S_j = s_j) = e(x_j | s_j)$$

• the transition and emission probabilities do not depend on the position in the Markov chain (do not depend on the index ) j

$$p(x_1 \dots x_m, s_1 \dots s_m) = t(s_1) \prod_{j=2}^m t(s_j | s_{j-1}) \prod_{j=1}^m e(x_j | s_j)$$

- Initial state parameters t(s) for  $s \in \{1, 2, ..., k\}$
- Transition parameters t(s'|s) for  $s, s' \in \{1, 2, ..., k\}$

### **HMM Learning**

• We collect a fully observed dataset  $\{X_i, S_i\}_{i=1}^N$ 

#### Training set:

- 1 Pierre/NNP Vinken/NNP ,/, 61/CD years/NNS old/JJ ,/ join/VB the/DT board/NN as/IN a/DT nonexecutive/JJ di Nov./NNP 29/CD ./.
- 2 Mr./NNP Vinken/NNP is/VBZ chairman/NN of/IN Else N.V./NNP ,/, the/DT Dutch/NNP publishing/VBG group/3 Rudolph/NNP Agnew/NNP ,/, 55/CD years/NNS old/JJ chairman/NN of/IN Consolidated/NNP Gold/NNP Fields/N,/, was/VBD named/VBN a/DT nonexecutive/JJ director/this/DT British/JJ industrial/JJ conglomerate/NN ./.

38,219 lt/PRP is/VBZ also/RB pulling/VBG 20/CD peopl of/IN Puerto/NNP Rico/NNP ./, who/WP were/VBD help Huricane/NNP Hugo/NNP victims/NNS ./, and/CC sendin them/PRP to/TO San/NNP Francisco/NNP instead/RB ./

#### Maximum Likelihood Estimate:

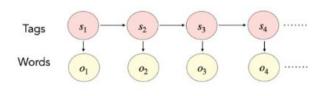
$$\max_{t(\cdot|\cdot),e(\cdot|\cdot)} \prod_{i=1}^{N} P(X_i, S_i)$$

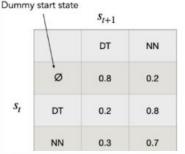
$$t(s'|s) = \frac{\text{count}(s \to s')}{\text{count}(s)}$$

$$e(x \mid s) = \frac{\text{count}(s \to x)}{\text{count}(s)}$$

## **HMM Example**

### Sequence probability





	$o_t$		
	the	cat	
DT	0.9	0.1	
NN	0.5	0.5	

What is the joint probability P(the cat, DT NN)?

- A) (0.8 \* 0.8) \* (0.9 \* 0.5) B) (0.2 \* 0.8) \* (0.9 \* 0.5)
- C) (0.3 \* 0.7) \* (0.5 \* 0.5)

## **HMM Learning Example**

- 1. the/DT cat/NN sat/VBD on/IN the/DT mat/NN
- 2. Princeton/NNP is/VBZ in/IN New/NNP Jersey/NNP
- 3. the/DT old/NN man/VB the/DT boats/NNS

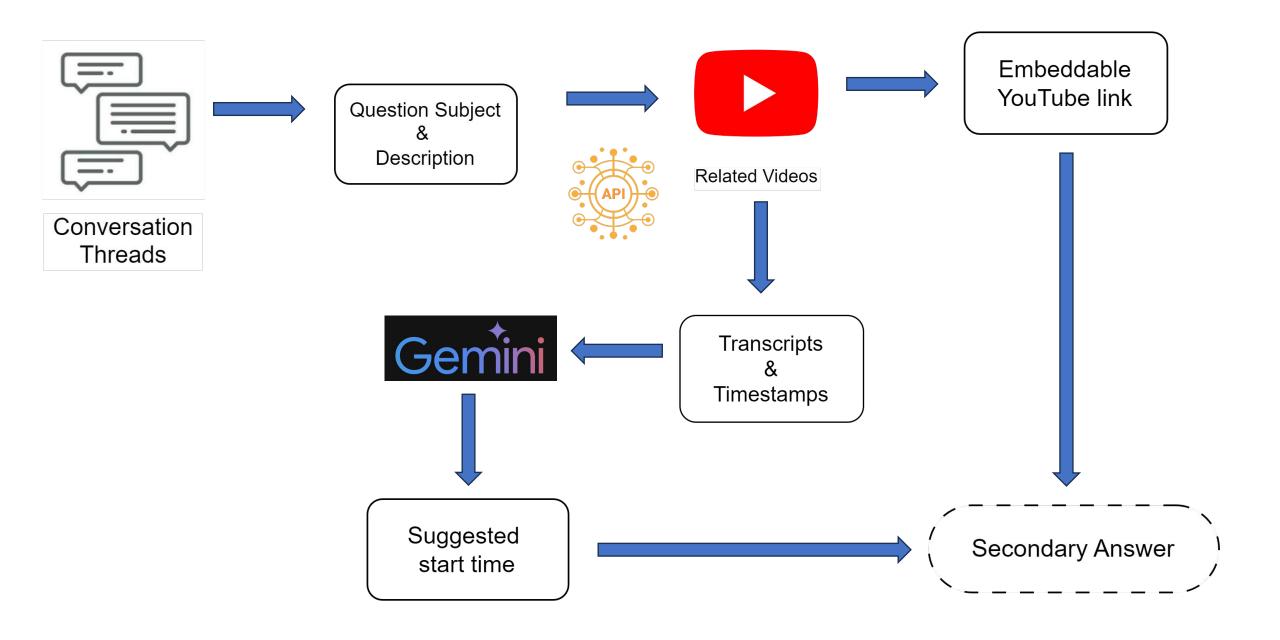
$$t(\mathbf{NN} | \mathbf{DT}) = \frac{3}{4}$$
  
 $e(\mathbf{cat} | \mathbf{NN}) = \frac{1}{3}$ 

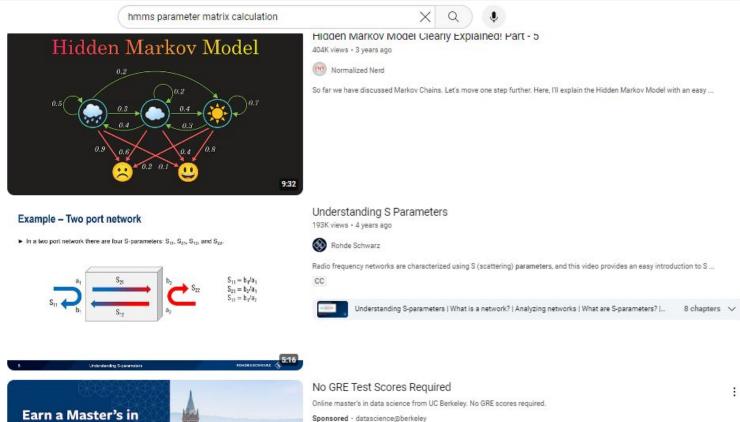
#### Maximum Likehood Estimate:

$$\max_{t(\cdot|\cdot),e(\cdot|\cdot)} \prod_{i=1}^{n} P(X_i, S_i)$$

$$t(s'|s) = \frac{\text{count}(s \to s')}{\text{count}(s)}$$

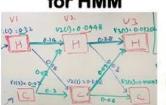
$$e(x \mid s) = \frac{\text{count}(s \to x)}{\text{count}(s)}$$













#### Viterbi Algorithm | HMM | Solved Decoding Example

90K views • 4 years ago

Binod Suman Academy

How to solve Hidden Markov Model Decoding problem. Markov Model Introduction: https://youtu.be/onSi24IM47U Markov Model ...

# HMMs parameter matrix calculation

How can one compute the emission and transition matrices for a Hidden Markov Model (HMM) based on a 3-character sequence? Additionally, could you provide the relevant formulas for these calculations?

