

NLP - Assignment 2

Pos-tagging and Viterbi 1

Initialization - I פל

המילה הראשונה: $w_1 = \text{"fish"}$, $t=1$

הסתברות C לפי הנוסחה: $C(i, 1) = \pi(i) \cdot B(i, w_1)$

$$C(N, 1) = 0.6 \cdot 0.7 = 0.42 \quad \text{פיקור N}$$

$$C(V, 1) = 0.4 \cdot 0.1 = 0.04 \quad \text{פיקור V}$$

הסתברות D המילה הראשונה: $D(N, 1) = 0$

$$D(N, 1) = 0$$

$$D(V, 1) = 0$$

Forward Pass II פל

המילה השנייה: $w_2 = \text{"swim"}$, $t=2$

הסתברות C לפי הנוסחה: $C(i, j) = \max_k [C(k, j-1) \cdot A(k, i) \cdot B(i, w_j)]$

הסתברות $C(N, 2)$ (שור, N , swim):

בדוק C הא (השור, N , swim)

$$0.42 \cdot 0.2 \cdot 0.1 = 0.0084 \quad \text{1. מסתכל N}$$

$$0.04 \cdot 0.5 \cdot 0.1 = 0.0020 \quad \text{2. מסתכל V}$$

\Rightarrow הסתברות המקסימלית היא 0.0084 והיא הסתברות $C(N, 2)$

המילה השלישית (argmax) N - זוהי הסתברות $D(N, 2)$

הסתברות $C(V, 2)$ (שור, V , swim):

בדוק C הא (השור, V , swim)

$$0.42 \cdot 0.8 \cdot 0.4 = 0.1344 \quad \text{1. מסתכל N}$$

$$0.04 \cdot 0.5 \cdot 0.4 = 0.0080 \quad \text{2. מסתכל V}$$

\Rightarrow הסתברות המקסימלית היא 0.1344 והיא הסתברות $C(N, 2)$

המילה השלישית (argmax) N - זוהי הסתברות $D(N, 2)$

- The Viterbi decoding matrix C

	$w_1(\text{fish})$	$w_2(\text{swim})$
N	0.42	0.0084
V	0.04	0.1344

Matrix D

	$w_1(\text{fish})$	$w_2(\text{swim})$
N	0	1 (N)
V	0	1 (N)

* המרחק: פחותה מ-1

1 מרחק N

- The final most-likely tag sequence

Backward Pass - המרחק, המרחק, המרחק

1. סוף המרחק (t=2) - המרחק, המרחק, המרחק C.

המרחק, המרחק, המרחק 0.1344, המרחק, המרחק V.

$$\text{Tag}_2 = V \quad \leftarrow$$

2. סוף המרחק (t=1) - המרחק, המרחק, המרחק (V, 2)

המרחק, המרחק, המרחק, המרחק, המרחק V, המרחק, המרחק N.

$$\text{Tag}_1 = N \quad \leftarrow$$

המרחק, המרחק: $N \rightarrow V$