(1)

(b)

$p(\text{I see you running around})$

$\Rightarrow$ We know, $p(w_1, w_2, w_3) = p(w_1 \mid w_2, w_3) \cdot p(w_2 \mid w_1, w_3) \cdot p(w_3 \mid w_1, w_2)$

$\therefore p(\text{I see you running around})$

$= p(\text{around} \mid \text{see you running}) \cdot p(\text{running} \mid \text{I see you})$

$\cdot p(\text{you} \mid \langle\text{start}\rangle\ \text{I see}) \cdot p(\text{see} \mid \langle\text{start}\rangle\ \langle\text{start}\rangle\ \text{I})$

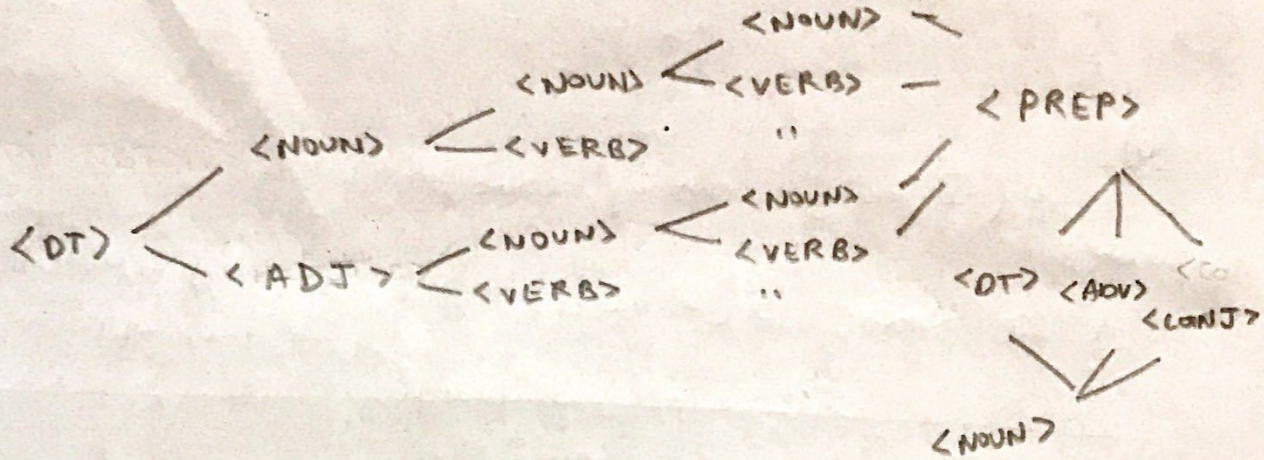
$\cdot p(\text{I} \mid \langle\text{start}\rangle\ \langle\text{start}\rangle\ \langle\text{start}\rangle)$

(c)

N-gram models have higher computational costs to gain better accuracy (larger N). Therefore, since, language models usually require large N, N-gram models tend to have higher costs and are insufficient.

(2)

(a)



∴ We can have 24 possible tag sequences.

(b)

1. $p(\text{can} \mid \text{NOUN}) = 0.001$ (Given) TRUE

2. $P(\text{VERB} \mid \text{AUX}) = 0.5$ (Given)

$\Rightarrow P(\text{not VERB} \mid \text{AUX}) = 1 - 0.5 = 0.5$ TRUE

3. $P(\text{can} \mid \text{AUX}) > P(\text{can} \mid \text{NOUN})$ (Given)

$\Rightarrow P(\text{she} \mid \text{PRON}\ \text{can} \mid \text{AUX}) > P(\text{she} \mid \text{PRON}\ \text{can} \mid \text{NOUN})$

TRUE

∴ Answer : E (1, 2 and 3)

4

A

(b) No, one hot encoding cannot be used for cosine similarity since they also represent a certain word given a vocabulary and when we measure the similarity using cosine rule, we end up getting 0 for every comparison.

Therefore, one-hot encoding does not place similar words closer to each other in vector space.

B

(c) $V_{boy} - V_{brother} \simeq V_{girl} - V_{sister}$

Since, boy and girl are similar to king and queen and man and women are similar to brother and sister (they all have opposite genders)

∴ Answer : C

①

(a) Language Models are used in parts-of-speech tagging and handwriting recognition.

(3)

(b)

a) deer-elk and d) okapi-caribou are the most similar pairs since both need 1 link to reach from the first word to the second.

Answer: a and d

(4) (A)

(a)

have = $[0,0,0,0,1]^T$

a = $[1,0,0,0,0]^T$

good = $[0,0,1,0,0]^T$

great = $[0,0,0,1,0]^T$

day = $[0,1,0,0,0]^T$

(2)

(c)

(a) $P(\text{planes} \mid \text{VERB}) = \frac{1}{3}$

[planes | VERB appear once out of all the 3 planes words]

(b) $P(\text{planes} \mid \text{NOUN}) = \frac{2}{3}$

[planes | NOUN appear twice out of all the 3 planes words]

(3)

(a)

(i) synonym

(ii) antonym

(iii) hyponym

(iv) meronym

(v) holonym