

4 | תרגיל NLP (67658)

יואב שפירא | ת"ז: 312492838 & נדב פוקס | ת"ז: 206073108

2 בינואר 2023

שאלה 1

1. מימוש

2. מימוש

3. מימוש

4. *Evaluation*: מדדנו את ה-*accuracy* עבור משפט, לפי: $\frac{|T \cap T'|}{|T|}$, כש T הוא העץ האמיתי - ה-*golden standard*, ו- T' הוא העץ שאנחנו מצאנו בעזרת האלגוריתם. עבור כל ה-*test set* קיבלנו בממוצע דיוק של: 0.38.

5. הוספנו את הפיצ'רים הבאים: לכל קשת פוטנציאלית (u, v) בדקנו מה התגיות של המילים שלפניהן ושל אחריהן (מילה אחת בכל כיוון), ובדקנו את כמות המילים ביניהן (עד ל-30 מילים). אכן הביצועים השתפרו משמעותית וקיבלנו ממוצע של 0.696.

שאלה 2

1. *Arc - Standard*: בסה"כ 14 מעברים:

(א) *SHIFT*

(ב) *SHIFT*

(ג) *Left - Arc*

(ד) *SHIFT*

(ה) *SHIFT*

(ו) *SHIFT*

(ז) *Left - Arc*

Right – Arc (ח)

SHIFT (ט)

SHIFT (י)

Left – Arc (יא)

Right – Arc (יב)

Right – Arc (יג)

Left – Arc (יד)

2. *Arc – Eager*: סה"כ 11 מעברים:

SHIFT (א)

Left – Arc (ב)

Right – Arc (ג)

Right – Arc (ד)

SHIFT (ה)

Left – Arc (ו)

Right – Arc (ז)

Reduce (ח)

SHIFT (ט)

Left – Arc (י)

Right – Arc (יא)