

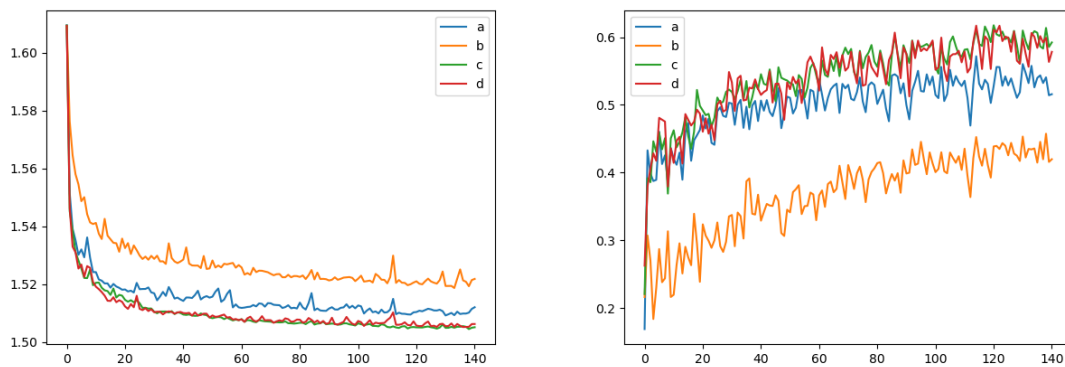
הפרמטרים היו:

Dropout של 0.1, ששומש בשתי השכבות של ה-LSTM. גודל הזכרון של ה-LSTM היה 50. גודל הזכרון של ה-char embedding היה 35, וגודל ה-embedding של התווים היה 35. גודל ה-embedding של המילים היה 50, והשתמשנו ב-pretrained embedding שניתנו.

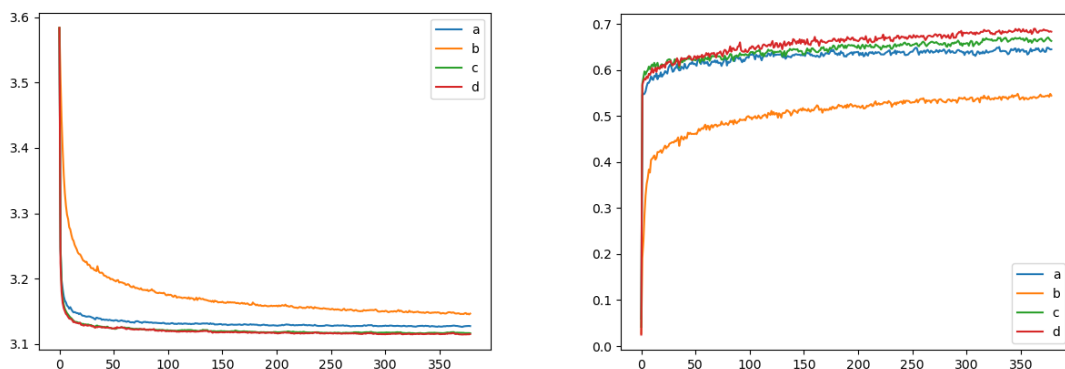
מאחר והיה נראה שגם עם batching הביצועים לא יהיו מוצלחים מאוד, בעיקר בגלל צורך בריפוד מילים (עבור סעיפים b,d) ומשפטים (עבור כל הסעיפים) והתעלמות לאחר מכן, החלטנו לא לבצע batching.

כמו בתרגיל הקודם, החלפנו בהסתברות מסוימת מילים בתג שמסמן UNKOWN באימון. אם הוחלט להחליף מילה, הייתה הסתברות להחליף גם את ה-prefix/suffix שלה ב-UNKOWN.

הגרפים: עבור ner:



עבור pos:



אני אוהב ש-a,b הם הכי גרועים, אבל כשמחברים אותם ביחד (d) הם הכי טובים :