

מגישים: אור שחר, 209493709, יהונתן שאקי 204920367.

חילקנו 70-30 בין אימון לטסט.

תוך כמות קטנה יחסית של מעברים על כל האימון (בערך 8 עד 15) הגענו לדיוק של מאה או 99% על הטסט (והיו כמה פעמים אפילו מאה אחוז לגמרי). מבחינת זמן, לקח ממש דקה או שתיים עד להגעה לדיוק הזה.

בעיה שניתקלנו בה: במקור, לפעמים התהליך היה קורה כמו שמתואר לעיל ולפעמים הרשת הייתה מגיעה לאיזה מינימום מקומי עם תחזית קבועה של 0 או 1 (קצת מוזר כי ה-data מאוד מאוזן). החלטנו שברגע שהרשת מגיעה לתחזית מוגזמת (0.85 אחוז ומעלה) לאחת מהמחלקות, משנים את ה-loss כך שיהיה ממושקל ויתן פחות משקל למחלקה אותה חוזים רוב הזמן (וחוזרים למצב הרגיל כשהרשת מפסיקה לעשות את זה). באופן כזה, נראה שהבעיה כבר לא קורת. מה שכן, בסיטואציות בהם זה נדרש לוקח לרשת מעט יותר זמן להתאמן (עד 30~ אפוקים).

השמשנו ברשת עם שתי שכבות של LSTM, גודל זכרון 128, dropout של 0.5 ו-learning rate של 0.001. כמובן, הפלט של ה-LSTM עבר לשכבת sigmoid (כולל ליניארית לפני).

הדוגמאות שיצרנו נמצאות בתקיה data/part\_one, ביחד עם gen\_examples.py.