**תרגיל בית 3 - רשתות RNN**

# הקדמה

במסגרת תרגיל בית זה מימשנו מספר רשתות RNN ליצר פסקאות בעלות משמעות/היגיון, כאשר הרשת אומנה על מאגר Penn Treebank. הרשתות השונות אומנו במספר קונפיגורציות, כאשר בסופו של דבר נבחרה הרשת שהניבה את הביצועים הטובים ביותר – 97.8 Preplexity. בנוסף לכך בוצעה השוואה לבחינת השפעת העיבוד המקדים (augmentation) על טיב התוצאות.

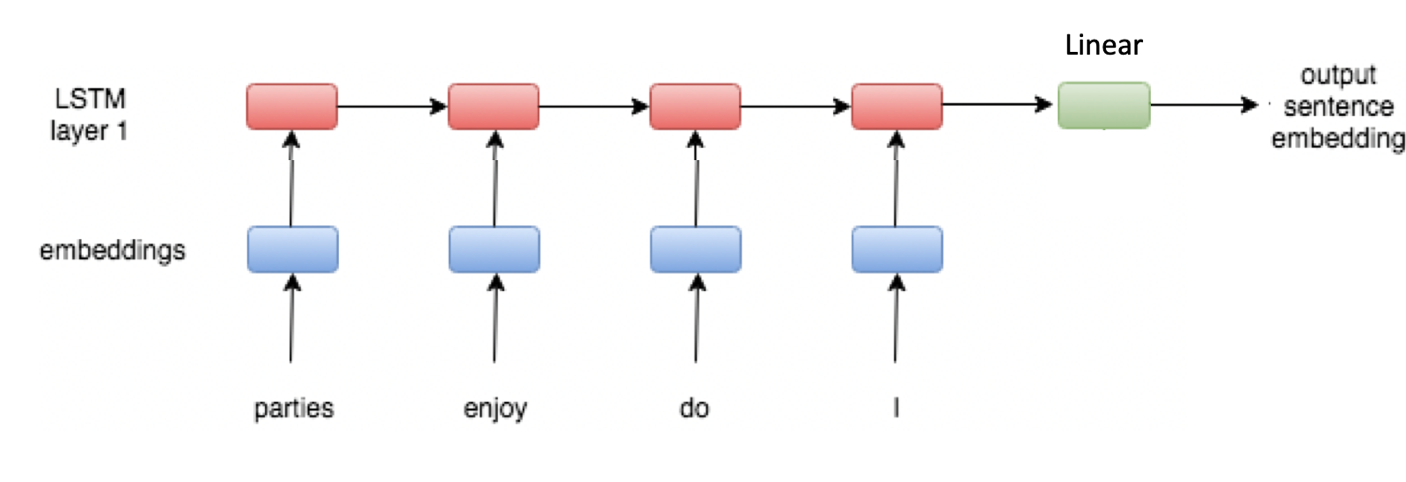
# תהליך הלמידה

לצורך בחירת המודל המיטבי ניסינו מספר רשתות, כאשר בסופו של דבר הוחלט על רשת אחת שנתנה את הביצועים הטובים ביותר. פונקציית ההפסד שהוגדרה היא NLLL (שיחד עם Soft-max מהווה הפסד Cross-entropy) ואלגוריתם ADAM שימש לצורך אופטימיזציה. בנוסף עשינו שימוש בקצב לימוד משתנה התלוי באם ישנו שיפור במשך 10 epoch, לצורך כך השתמשנו במתזמן ReduceLROnPlateau של Pytorch.

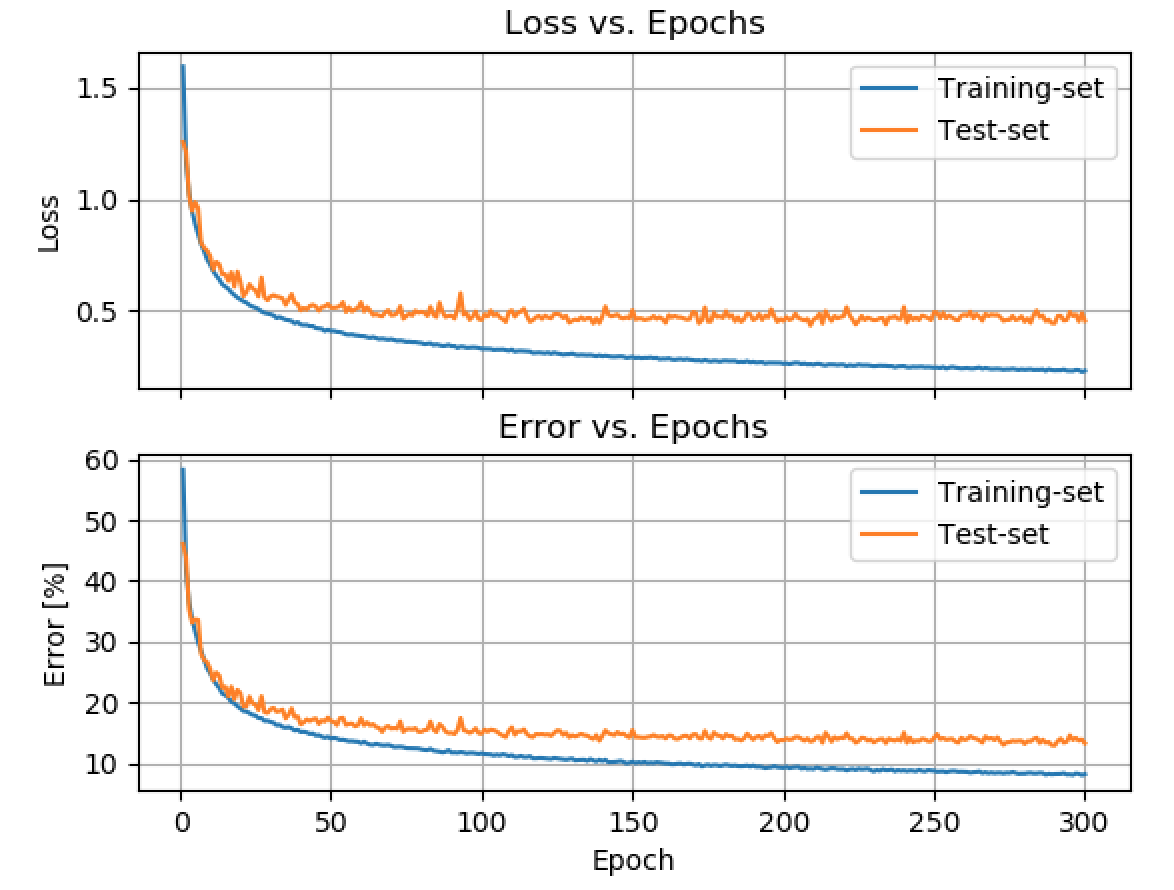
המודל הנבחר:

* מודל (LSTM) - רשת בעלת שכבה אחת של Embedding ומעלייה שכבה אחת של LSTM ולבסוף שכבה לינארית בגודל 249 ל-10,000 (שהוא מספר המילים במילון). גודל שתי השכבות הוא 249 פרטים. לטובת חיסכון במספר הפרמטרים, השווינו את המשקלים בין השכבות. בנוסף הוספנו גם Dropout גם ב-LSTM וגם לאחריו במעבר לשכבה הלינארית. לרשת סה"כ 2,998,000 משקלים. ראה איור 1.

|  |
| --- |
|  |



איור 1 – מודל 1



*איור 2 - הפסדים ושגיאות בזמן הלמידה של המודל המיטבי*

# סיכום ומסקנות

התחלנו עם ניסיון להתמודד עם הבעיה בעזרת מודלים המוכוונים לרעיון עיבוד התמונה ע״י דחיסת התמונה ומיצוע המידע בה ע״י מספר גדול של פילטרים על פני מימד קטן, כמו מודל הVGG. אך התוצאות של המודלים הללו לא סיפקו את דרישתנו בשל הגבלת מספר הפרמטרים. לאחר מכן החלטנו לעבור למודל של DenseNet שמשתמש בעקרון דחיסת שכבות הקונבולוציה עם Append מעל השכבות. המודל החדש אפשר לנו להוסיף יותר שכבות על חשבון מספר פילטרים קטן יותר בכל שכבה. מודל זה הקטין משמעותית השגיאה על פני סט הבוחן ב-6.45%. ההקטנה אכן משמעותית אם אנחנו לוקחים בחשבון את הגבלת מספר הפרמטרים (50k).

לאחר שבחרנו במודל התחלנו בניסיון למצות את המידע בתמונות על ידי שינוי של התמונות עם שינוי בהירות, היפוך או ניגודיות של התמונות. השינויים הללו בנוסף לאפשרות להתבונן אחרת על התמונות, הרחיבו את כמות התמונות השונות שהמודל אומן עליהם, מכיוון שכל Batch התמונות נדגמות עם הסתברות לשינוי כל תמונה לאחד האפשרויות (עם אפשרות להפעלה של מספר שינויים יחדיו). לכן ציפינו שדווקא אפשרות ד' תביא לשיפור בתוצאות, אך בפועל דווקא האופציה הפשוטה ביותר - א', הניבה את התוצאות הטובות ביותר.