

הערה לתרגיל: • מסמן מכפלה סקלרית או וקטורית רגילה (לפי ההקשר), \otimes מסמן מכפלה חיצונית.

שאלה 1

א. במקרה בו $i = j$:

$$\begin{aligned}\nabla_{\theta_i} \text{softmax}(\theta)_i &= \frac{\exp(\theta_i) \cdot \sum_j \exp(\theta_j) - \exp(\theta_i) \cdot \exp(\theta_i)}{(\sum_j \exp(\theta_j))^2} \\ &= \text{softmax}(\theta)_i - (\text{softmax}(\theta)_i)^2 = \text{softmax}(\theta)_i (1 - \text{softmax}(\theta)_i)\end{aligned}$$

במקרה בו $i \neq j$:

$$\begin{aligned}\nabla_{\theta_i} \text{softmax}(\theta)_j &= \frac{0 - \exp(\theta_j) \cdot \nabla_{\theta_i} \sum_k \exp(\theta_k)}{(\sum_k \exp(\theta_k))^2} = -\frac{\exp(\theta_j) \cdot \exp(\theta_i)}{(\sum_k \exp(\theta_k))^2} \\ &= -\text{softmax}(\theta)_i \cdot \text{softmax}(\theta)_j = \text{softmax}(\theta)_i (0 - \text{softmax}(\theta)_j)\end{aligned}$$

סה"כ נקבל:

$$\nabla_{\theta_i} \text{softmax}(\theta)_j = \text{softmax}(\theta)_i \cdot (\mathbb{I}\{i = j\} - \text{softmax}(\theta)_j) \Rightarrow$$

$$\nabla_{\theta_i} \text{softmax}(\theta) = \text{softmax}(\theta)_i \cdot (e_i - \text{softmax}(\theta)) \Rightarrow$$

$$\nabla_{\theta} \text{softmax}(\theta) = \text{Diag}(\text{softmax}(\theta)) - \text{softmax}(\theta) \otimes \text{softmax}(\theta)$$

כעת נגזור כל איבר בסכום $CE(y, \hat{y}) = -\sum_i y_i \cdot \log(\hat{y}_i)$ בנפרד עם חלוקה למקרים.

כאשר $i = j$:

$$\begin{aligned}\nabla_{\theta_i} - y_i \cdot \log(\hat{y}_i) &= \nabla_{\theta_i} - y_i \cdot \log(\text{softmax}(\theta)_i) = -\frac{y_i}{\text{softmax}(\theta)_i} \cdot \nabla_{\theta_i} (\text{softmax}(\theta)_i) \\ &= -\frac{y_i}{\text{softmax}(\theta)_i} \cdot (\text{softmax}(\theta)_i - \text{softmax}(\theta)_i^2) \\ &= -y_i(1 - \text{softmax}(\theta)_i) = y_i \text{softmax}(\theta)_i - y_i\end{aligned}$$

כאשר $i \neq j$:

$$\begin{aligned}\nabla_{\theta_i} - y_j \cdot \log(\hat{y}_j) &= \nabla_{\theta_i} - y_j \cdot \log(\text{softmax}(\theta)_j) = -\frac{y_j}{\text{softmax}(\theta)_j} \cdot \nabla_{\theta_i} (\text{softmax}(\theta)_j) \\ &= -\frac{y_j}{\text{softmax}(\theta)_j} \cdot (-\text{softmax}(\theta)_i \cdot \text{softmax}(\theta)_j) = y_j \cdot \text{softmax}(\theta)_i\end{aligned}$$

ובסה"כ קיבלנו:

$$\nabla_{\theta_i} CE(y, \hat{y}) = \sum_j (y_j \cdot softmax(\theta)_i) - y_i = softmax(\theta)_i - y_i$$

מכך ש- y הוא one-hot vector. ולכן:

$$\nabla_{\theta} CE(y, \hat{y}) = softmax(\theta) - y = \hat{y} - y$$

ב. נגדיר $z := xW_1 + b_1$ ונחשב:

$$\begin{aligned} \left(\frac{\partial z}{\partial x}\right)_{i,j} &= \frac{\partial z_i}{\partial x_j} = \frac{\partial}{\partial x_j} \sum_{k=1}^{D_x} x_k W_1[k, i] + b_{1,k} = \sum_{k=1}^{D_x} \frac{\partial}{\partial x_j} x_k W_1[k, i] + b_{1,k} = W_1[j, i] \\ \Rightarrow \frac{\partial z}{\partial x} &= W_1^T \end{aligned}$$

כאשר $W_1[i, j]$ מסמל את הערך בתא ה- (i, j) במטריצה W_1 .

כעת נגדיר $\theta := hW_2 + b_2$ ומסעיף א' נקבל:

$$\frac{\partial J}{\partial \theta} = \hat{y} - y$$

בנוסף נחשב בעזרת כלל השרשרת:

$$\frac{\partial \theta}{\partial z} = \frac{\partial \theta}{\partial h} \cdot \frac{\partial h}{\partial z} = \left(\frac{\partial}{\partial h} hW_2 + b_2 \right) \cdot \frac{\partial}{\partial z} \sigma(z) = W_2^T \cdot \text{Diag}(\sigma'(z))$$

כאשר הנגזרת הראשונה היא באופן זהה לחישוב של $\frac{\partial z}{\partial x}$, ואת הנגזרת של סיגמואיד חישבנו

בתרגיל הקודם. עכשיו נוכל להיעזר בכלל השרשרת ולחשב את הנגזרת הרצויה:

$$\frac{\partial J}{\partial x} = \frac{\partial J}{\partial \theta} \cdot \frac{\partial \theta}{\partial z} \cdot \frac{\partial z}{\partial x} = (\hat{y} - y) \cdot (W_2^T \cdot \text{Diag}(\sigma'(z))) \cdot W_1^T$$

ג. נחשב את הנגזרות הנדרשות למימוש:

$$\frac{\partial J}{\partial W_1} = \frac{\partial J}{\partial \theta} \cdot \frac{\partial \theta}{\partial z} \cdot \frac{\partial z}{\partial W_1} = x^T \otimes ((\hat{y} - y) \cdot (W_2^T \cdot \text{Diag}(\sigma'(z))))$$

$$\frac{\partial J}{\partial b_1} = \frac{\partial J}{\partial W_1} = \frac{\partial J}{\partial \theta} \cdot \frac{\partial \theta}{\partial z} \cdot \frac{\partial z}{\partial b_1} = (\hat{y} - y) \cdot (W_2^T \cdot \text{Diag}(\sigma'(z)))$$

$$\frac{\partial J}{\partial W_2} = \frac{\partial J}{\partial \theta} \cdot \frac{\partial \theta}{\partial W_2} = h \otimes (\hat{y} - y)$$

$$\frac{\partial J}{\partial b_2} = \frac{\partial J}{\partial \theta} \cdot \frac{\partial \theta}{\partial b_2} = (\hat{y} - y)$$

שאלה 2

א. נגדיר $\theta = h^{(t)}U + b_2, z = h^{(t-1)}H + e^{(t)}I + b_1$ ואז נחשב לפי כלל השרשרת:

i.

$$\frac{\partial J^{(t)}}{\partial U} = \frac{\partial J^{(t)}}{\partial \theta} \cdot \frac{\partial \theta}{\partial U} = h^{(t)T} \otimes (\hat{y}^{(t)} - y^{(t)})$$

ii.

$$\frac{\partial J^{(t)}}{\partial b_2} = \frac{\partial J^{(t)}}{\partial \theta} \cdot \frac{\partial \theta}{\partial b_2} = \hat{y}^{(t)} - y^{(t)}$$

iii.

$$\frac{\partial J^{(t)}}{\partial L_{x^{(t)}}} = \frac{\partial J^{(t)}}{\partial e^{(t)}} = \frac{\partial J^{(t)}}{\partial \theta} \cdot \frac{\partial \theta}{\partial h^{(t)}} \cdot \frac{\partial h^{(t)}}{\partial e^{(t)}} = (\hat{y}^{(t)} - y^{(t)}) \cdot U^T \cdot \text{Diag}(h^{(t)} - (h^{(t)})^2) \cdot I^T$$

כאשר השתמשנו בנגזרת של סיגמואיד, אותה חישבנו בתרגיל הקודם.

iv.

$$\left. \frac{\partial J^{(t)}}{\partial I} \right|_{(t)} = \left. \frac{\partial J^{(t)}}{\partial z} \cdot \frac{\partial z}{\partial I} \right|_{(t)} = (e^{(t)})^T \cdot \frac{\partial J^{(t)}}{\partial z}$$

נחשב:

$$\frac{\partial J^{(t)}}{\partial z} = \frac{\partial J^{(t)}}{\partial \theta} \cdot \frac{\partial \theta}{\partial h^{(t)}} \cdot \frac{\partial h^{(t)}}{\partial z} = (\hat{y}^{(t)} - y^{(t)}) \cdot U^T \cdot \frac{\partial \sigma(z)}{\partial z}$$

נחשב:

$$\frac{\partial \sigma(z)}{\partial z} = \text{diag}(h^{(t)} - h^{(t)2})$$

נציב חזרה:

$$\frac{\partial J^{(t)}}{\partial z} = (\hat{y}^{(t)} - y^{(t)}) \cdot U^T \cdot \frac{\partial \sigma(z)}{\partial z} = (\hat{y}^{(t)} - y^{(t)}) \cdot U^T \cdot \text{diag} \left(h^{(t)} - h^{(t)^2} \right)$$

נציב חזרה:

$$\left. \frac{\partial J^{(t)}}{\partial I} \right|_{(t)} = \frac{\partial J^{(t)}}{\partial z} \cdot (e^{(t)})^T = (e^{(t)})^T \cdot (\hat{y}^{(t)} - y^{(t)}) \cdot U^T \cdot \text{diag} \left(h^{(t)} - h^{(t)^2} \right)$$

.v

$$\left. \frac{\partial J^{(t)}}{\partial H} \right|_{(t)} = \frac{\partial J^{(t)}}{\partial z} \cdot \frac{\partial z}{\partial H} \Big|_{(t)} = h^{(t-1)T} \cdot (\hat{y}^{(t)} - y^{(t)}) \cdot U^T \cdot \text{diag} \left(h^{(t)} - h^{(t)^2} \right)$$

.vi

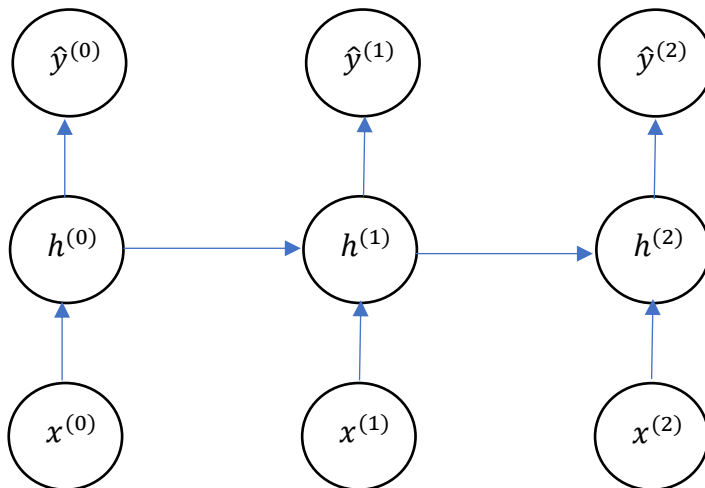
$$\frac{\partial J^{(t)}}{\partial b_1} = \frac{\partial J^{(t)}}{\partial z} \cdot \frac{\partial z}{\partial b_1} = (\hat{y}^{(t)} - y^{(t)}) \cdot U^T \cdot \text{diag} \left(h^{(t)} - h^{(t)^2} \right)$$

כאשר חישובי הביניים זהים לשאלה 1.

.vii

$$\frac{\partial J^{(t)}}{\partial h^{(t-1)}} = \frac{\partial J^{(t)}}{\partial z} \cdot \frac{\partial z}{\partial h^{(t-1)}} = (\hat{y}^{(t)} - y^{(t)}) \cdot U^T \cdot \text{diag} \left(h^{(t)} - h^{(t)^2} \right) \cdot H^T$$

.ב



$$z^{(t)} = h^{(t-1)}H + e^{(t)}I + b_1$$

נזכיר כי חישובנו $\frac{\partial J^{(t)}}{\partial h^{(t-1)}} = (\hat{y}^{(t)} - y^{(t)}) \cdot U^T \cdot \text{diag}(h^{(t)} - h^{(t)^2}) \cdot H^T$ מעתה לא נכתוב את כל הביטוי.

i.

$$\frac{\partial h^{(t)}}{\partial e^{(t)}} = \text{Diag}(h^{(t)} - (h^{(t)})^2) \cdot I^T$$

$$\frac{\partial J^{(t)}}{\partial L_{x^{(t-1)}}} = \frac{\partial J^{(t)}}{\partial e^{(t-1)}} = \frac{\partial J^{(t)}}{\partial h^{(t-1)}} \cdot \frac{\partial h^{(t-1)}}{\partial e^{(t-1)}} = \frac{\partial J^{(t)}}{\partial h^{(t-1)}} \cdot \text{Diag}(h^{(t-1)} - (h^{(t-1)})^2) \cdot I^T$$

ii.

$$h^{(t-1)} = \sigma(z^{(t-1)}) \text{, כמו כן, } \frac{\partial \sigma(z^{(t)})}{\partial z^{(t)}} = \text{Diag}(h^{(t)} - h^{(t)^2})$$

$$\begin{aligned} \left. \frac{\partial J^{(t)}}{\partial H} \right|_{(t-1)} &= \frac{\partial J^{(t)}}{\partial h^{(t-1)}} \cdot \frac{\partial \sigma(z^{(t-1)})}{\partial z^{(t-1)}} \cdot \left. \frac{\partial z^{(t-1)}}{\partial H} \right|_{(t-1)} \\ &= \frac{\partial J^{(t)}}{\partial h^{(t-1)}} \cdot \text{diag}(h^{(t-1)} - h^{(t-1)^2}) \cdot \left. \frac{\partial z^{(t-1)}}{\partial H} \right|_{(t-1)} \end{aligned}$$

נחשב:

$$\frac{\partial z^{(t-1)}}{\partial H} = \frac{\partial (h^{(t-2)}H + e^{(t-1)}I + b_1)}{\partial H} = h^{(t-2)}$$

נציב חזרה:

$$\left. \frac{\partial J^{(t)}}{\partial H} \right|_{(t-1)} = h^{(t-2)^T} \cdot \frac{\partial J^{(t)}}{\partial h^{(t-1)}} \cdot \text{diag}(h^{(t-1)} - h^{(t-1)^2})$$

iii.

$$\begin{aligned} \left. \frac{\partial J^{(t)}}{\partial I} \right|_{(t-1)} &= \frac{\partial J^{(t)}}{\partial h^{(t-1)}} \cdot \frac{\partial \sigma(z^{(t-1)})}{\partial z^{(t-1)}} \cdot \left. \frac{\partial z^{(t-1)}}{\partial I} \right|_{(t-1)} \\ &= e^{(t-1)^T} \cdot \frac{\partial J^{(t)}}{\partial h^{(t-1)}} \cdot \text{diag}(h^{(t-1)} - h^{(t-1)^2}) \end{aligned}$$

$$\left. \frac{\partial J^{(t)}}{\partial b_1} \right|_{(t-1)} = \frac{\partial J^{(t)}}{\partial z^{(t-1)}} \cdot \left. \frac{\partial z^{(t-1)}}{\partial b_1} \right|_{(t-1)} = 1 \cdot \frac{\partial J^{(t)}}{\partial h^{(t-1)}} \cdot \text{diag} \left(h^{(t-1)} - h^{(t-1)^2} \right)$$

שאלה 3

בשפות כמו עברית וערבית, שבהן יש לעתים מספר מיליות המרכיבות את אותה מילה, עשוי להיות יתרון למודל שפה מבוסס אותיות. זאת משום שלכל מילית יש משמעות של מילה נפרדת ולכן למעשה נתייחס לאותה המילה עם וריאציות שונות כמילים נפרדות, לדוגמה, המילה "אני" - ההטיות "ואני", "כשאני", יהיו מיוצגות כל אחת ע"י טוקן אחד ולכן "יאבדו" את המשמעות הסמנטית לעומת חלוקת המילים ל- "ו", "כש", "אני".

כמו כן, נאבד גם קשרים בין מילים כמו "כאשר" ו-"כש", "בתוך" ו-"ב".

חלוקה של המודל לאותיות מבטלת גם את הסיכון להיתקל במילה לא ידועה משום שכל תו שנקרא הוא בוודאות במילון שהוגדר.

מהצד השני, חלוקה של המודל למילים מאפשרות יותר אקספרסיביות במובן שהמילון הרבה יותר גדול ועל כן יש יותר "קשרים" בין מילים. מחלוקה כזו נקבל קשר סמנטי וסינטקטי בין מילים לעומת אותיות שלרוב אין משמעות סינטקטית או סינטקטית לשכיחות שהן מופיעות יחדיו במשפט. במשפט "הלכתי לבנק ומשכתי כסף" נקבל ככה"נ תלות יותר גדולה בין המופעים של "לבנק" ו"כסף" מאשר בין המופעים של האות ב' והאות ס'.

שאלה 4

נסמן:

$$\exp(x) = e^x, \quad \exp2(x) = 2^x$$

נראה כי:

$$\begin{aligned} \exp2\left(-\frac{1}{M}\sum_{i=1}^M \log_2 p(s_i | s_1, \dots, s_{i-1})\right) &= \exp2\left(-\frac{\log_2(e)}{\log_2(e)} \frac{1}{M}\sum_{i=1}^M \log_2 p(s_i | s_1, \dots, s_{i-1})\right) \\ &= \left(2^{\log_2(e)}\right)^{-\frac{1}{M}\sum_{i=1}^M \frac{1}{\log_2(e)} \log_2 p(s_i | s_1, \dots, s_{i-1})} \\ &= \exp\left(-\frac{1}{M}\sum_{i=1}^M \log_e p(s_i | s_1, \dots, s_{i-1})\right) \\ &= \exp\left(-\frac{1}{M}\sum_{i=1}^M \ln p(s_i | s_1, \dots, s_{i-1})\right) \end{aligned}$$

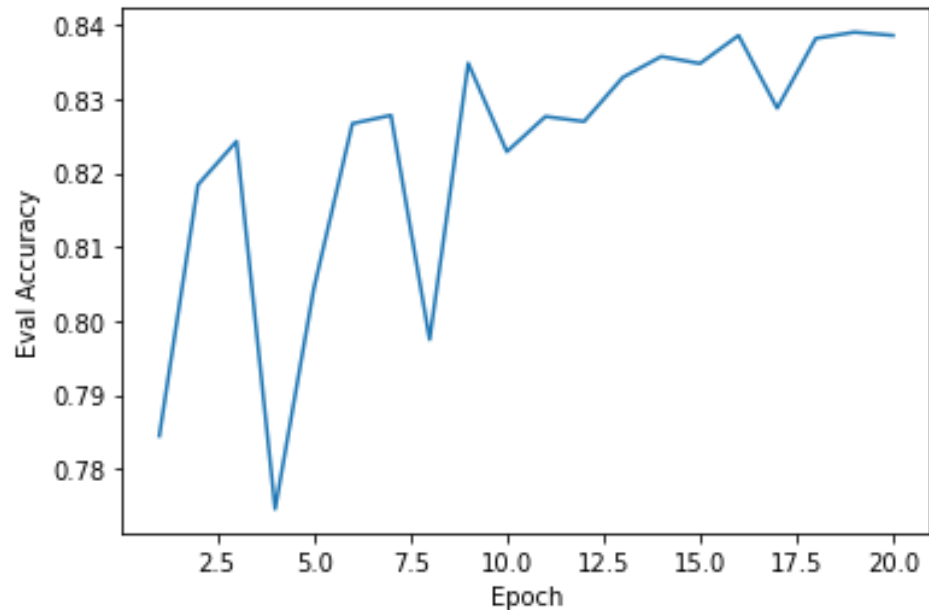
וקיבלנו את הדרוש. המעבר השני הוא מחוקי חזקות, השלישי הוא מחוקי לוגריתם והאחרון

הוא מכך ש- $\ln(x) := \log_e(x)$.

שאלה 5

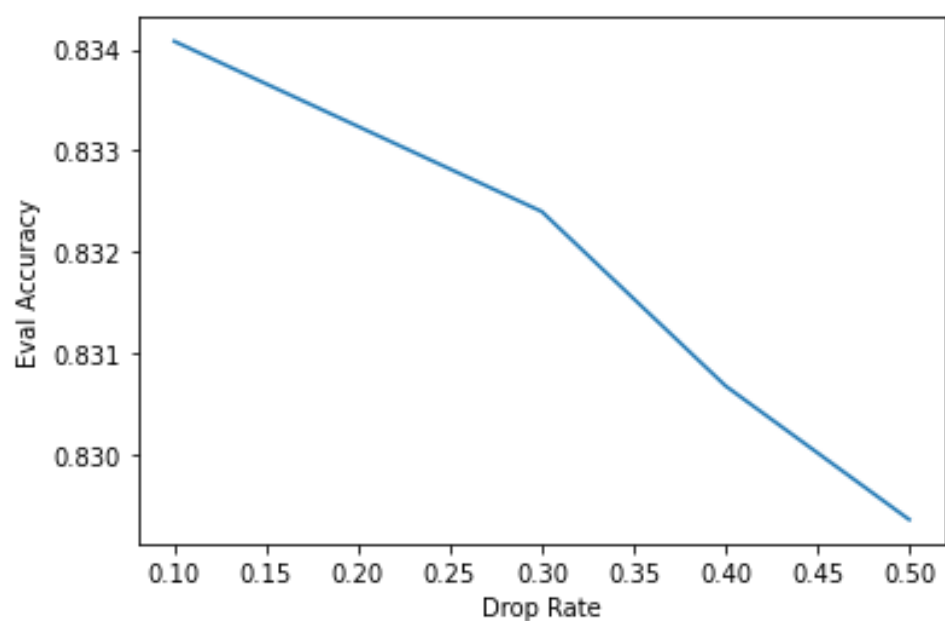
לכל הסעיפים, בחרנו את אותו רוחב לכל השכבות החביות – 300 פרמטרים (כמו פלט ה-*embedding*).

א. בסעיף זה מימשנו רשת עם 3 שכבות חביות, פונקציית אקטיבציה $LeakyReLU(0.1)$, 20 *epoch*-ים, גודל *batch* של 50 ו-*learning rate* התחלתי של 0.0005. להלן התוצאות:

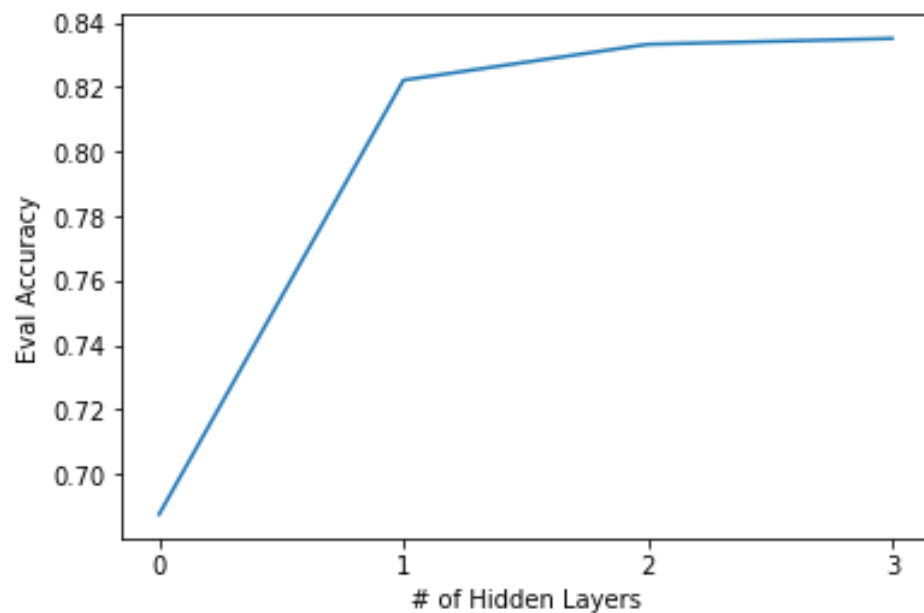


ניתן לראות שהמודל הגיע לדיוק גבוה מ-83.5%.

ב. הרצנו בסעיף זה רשת עם אותם פרמטרים כמו בסעיף הקודם. להלן התוצאות:

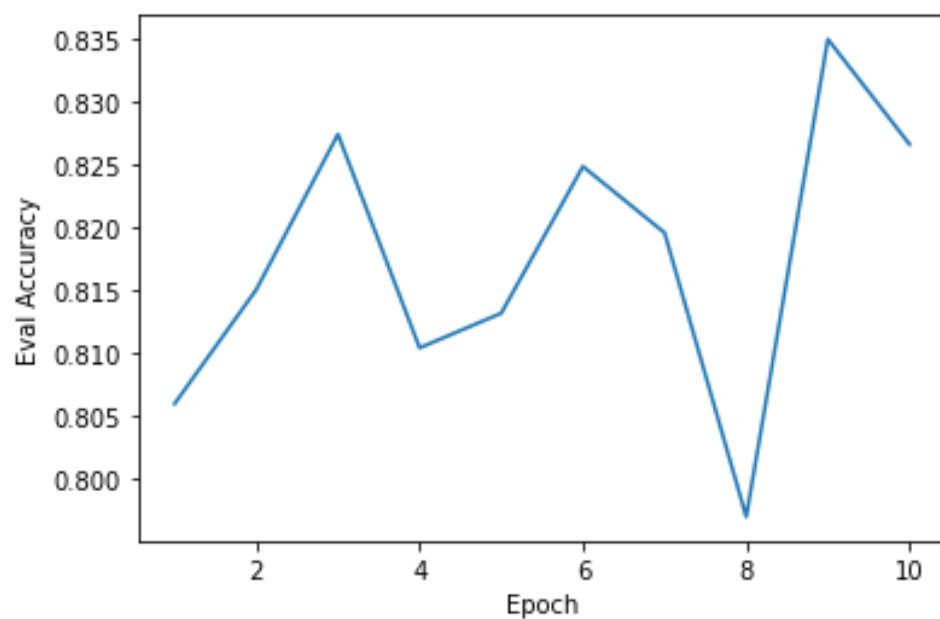


ג. השתמשנו באותם פרמטרים כמו סעיפים א' וב', פרט לכמות השכבות החביות. להלן התוצאות:

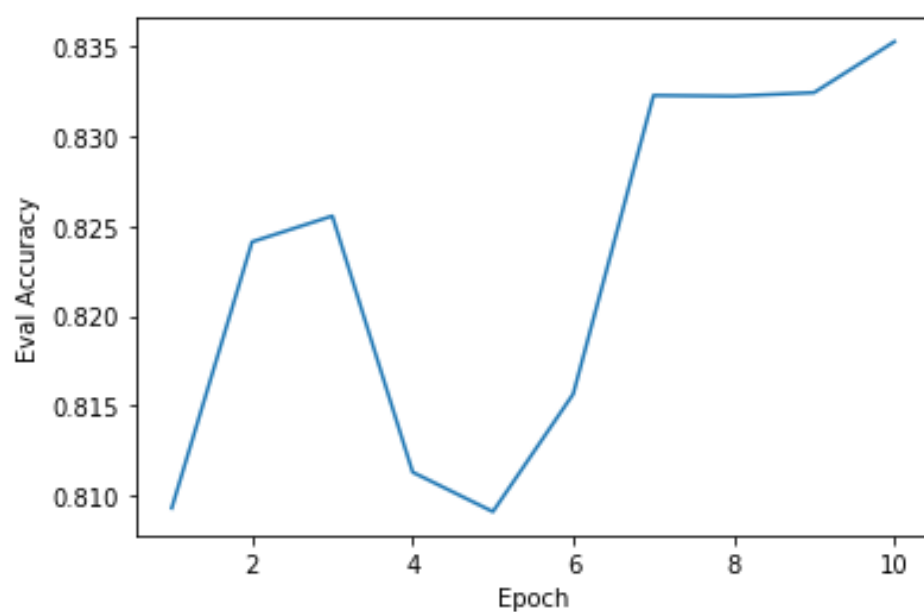


ניתן לראות שהדיוק עולה ככל שיש יותר שכבות חביות, ובפרט המודל הלינארי הוא עם הדיוק הנמוך ביותר. תוצאה זו הגיונית כיוון שהוספת שכבות רק מגדילה את האקספרסיביות של המודל, ומודל עם יותר שכבות דמויות יכול לדמות מודל קטן יותר (ע"י קיבוע השכבה הלינארית האחרונה למטריצת הזהות). תוצאה יותר מפתיעה היא שנראה שכבר מהשכבה השנייה והלאה נראה שהוספה של שכבות נוספות לא משפרת את הדיוק בהרבה (יתכן כי יש שינוי מגמה אם מוסיפים יותר מ-3 שכבות). נסיק כי רשת עם שתי שכבות חביות היא אקספרסיבית מספיק עבור המאגר הנתון.

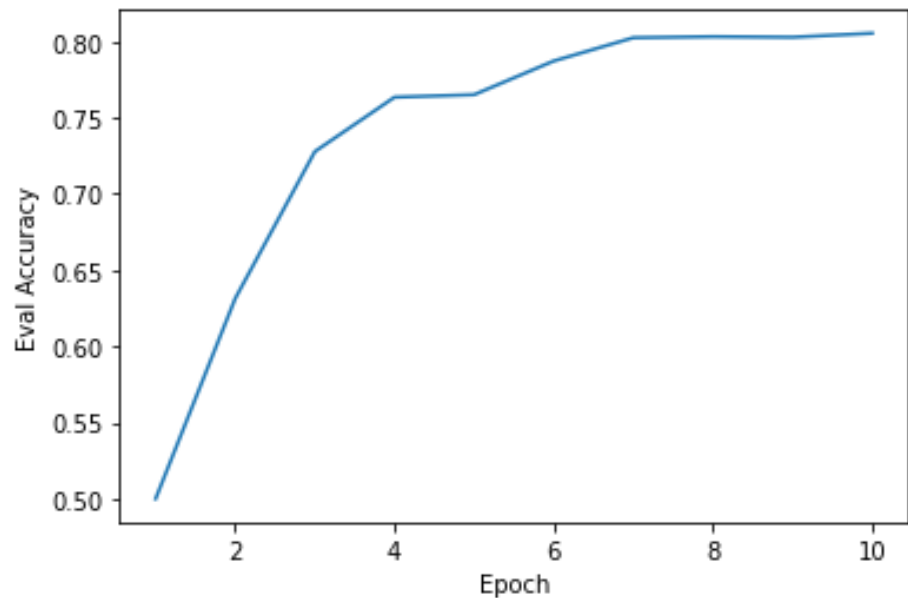
ד. בסעיף זה הרצנו פרמטרים זהים לסעיף א', אבל הגבלנו את הריצה ל-10 epochs-ים. הגרף עבור הפונקציה $ReLU$:



הגרף עבור הפונקציה $LeakyReLU(0.1)$:



הגרף עבור הפונקציה $Hardsigmoid$:



נראה כי הפונקציה *LeakyReLU* היא המוצלחת ביותר, מאחר והיא משיגה תוצאות יותר טובות משתי הפונקציות האחרות לאורך זמן וגם בסיום הריצה. בנוסף, נראה שפונקציית הסיגמואיד היא הכי פחות מוצלחת, ובפרט נראה שהדיוק שלה מתכנס בסביבות 80%, לעומת הפונקציות האחרות שלא מראות סימני התכנסות בעשרת ה-*epoch*-ים שהרצנו.

ה. חמש דוגמאות עליהן המודל הכי טוב שלנו (מסעיף א') טועה מצורפות בסוף הסעיף. בחמש הדוגמאות אפשר למצוא מילים בעלות סנטימנט מנוגד – בחלקן מוזכרת ציפיה חיובית לקראת הסרט (*I was looking forward to this movie. Trustworthy actors,*) *(interesting plot. Great atmosphere* ולאחר מכן אכזבה, בחלקן מובעים רגשות אמביוולנטיים (*Not bad. Not good. Passable 4*). לכן, הגיוני שהמודל לא הצליח לסווג אותן נכון – גם לבן-אדם תידרש כנראה קריאה מעמיקה של הביקורת כדי להבין את הסנטימנט הכללי שלה, ולא יספיק לעבור על המילים בנפרד ולקבוע אם הן חיוביות או שליליות.

1.

*"Worth the entertainment value of a rental, especially if you like action movies. This one features the usual car chases, fights with the great Van Damme kick style, shooting battles with the 40 shell load shotgun, and even terrorist style bombs. All of this is entertaining and competently handled but there is nothing that really blows you away if you've seen your share before.

The plot is made interesting by the inclusion*

of a rabbit, which is clever but hardly profound. Many of the characters are heavily stereotyped -- the angry veterans, the terrified illegal aliens, the crooked cops, the indifferent feds, the bitchy tough lady station head, the crooked politician, the fat federale who looks like he was typecast as the Mexican in a Hollywood movie from the 1940s. All passably acted but again nothing special.
I thought the main villains were pretty well done and fairly well acted. By the end of the movie you certainly knew who the good guys were and weren't. There was an emotional lift as the really bad ones got their just deserts. Very simplistic, but then you weren't expecting Hamlet, right? The only thing I found really annoying was the constant cuts to VDs daughter during the last fight scene.
Not bad. Not good. Passable 4."

הביקורת אמביוולנטית ומכילה גם מילים חיוביות וגם שליליות, וכנראה שלכן המודל טעה עליה.

2.

"Isaac Florentine has made some of the best western Martial Arts action movies ever produced. In particular US Seals 2, Cold Harvest, Special Forces and Undisputed 2 are all action classics. You can tell Isaac has a real passion for the genre and his films are always eventful, creative and sharp affairs, with some of the best fight sequences an action fan could hope for. In particular he has found a muse with Scott Adkins, as talented an actor and action performer as you could hope for. This is borne out with Special Forces and Undisputed 2, but unfortunately The Shepherd just doesn't live up to their abilities.
There is no doubt that JCVD looks better here fight-wise than he has done in years, especially in the fight he has (for pretty much no reason) in a prison cell, and in the final showdown with Scott, but look in his eyes. JCVD seems to be dead inside. There's nothing in his eyes at all. It's like he just doesn't care about anything throughout the whole film. And this is the leading man.
There are other dodgy aspects to the film, script-wise and visually, but the main problem is that you are utterly unable to empathise with the hero of the film. A genuine shame as I know we all wanted this film to be as special as it genuinely could have been. There are some good bits, mostly the

action scenes themselves. This film had a terrific director and action choreographer, and an awesome opponent for JCVD to face down. This could have been the one to bring the veteran action star back up to scratch in the balls-out action movie stakes.
Sincerely a shame that this didn't happen."

בדוגמא זו הביקורת על הסרט המדובר היא שלילית, אבל מוזכרים בה לחיוב סרטים קודמים, מה שכנראה גרם למודל לתת לה ציון סנטימנט חיובי ולטעות.

3.

'I first watched this movie back in the mid/late 80\'s, when I was a kid. We couldn\'t even get all the way through it. The dialog, the acting, everything about it was just beyond lame.
Here are a few examples... imagine these spoken real dramatically, way over-acted: "Oreegon? You\'re going to Oreegon? Why would anyone want to go to Oreegon?"
"Survivalists? Nobody ever told us about any survivalists!"
This movie was SO bad, my sister and I rented it again for her 16th birthday party, just so our friends could sit around and laugh at how awful it was. I don\'t think we were able to finish it then either!'

הביקורת בדוגמא זו היא לכאורה רעה – נאמר במפורש " *This movie was SO bad* ", אך הכותב/ת גם מציין/ת הנאה ממנו – צפייה עם חברים וצחוק, מה שכנראה גרם למודל לטעות.

4.

"I was looking forward to this movie. Trustworthy actors, interesting plot. Great atmosphere then ????? IF you are going to attempt something that is meant to encapsulate the meaning of life. First. Know it. OK I did not expect the directors or writers to actually know the meaning but I thought they may have offered crumbs to peck at and treats to add fuel to the fire-Which! they almost did. Things I didn't get. A woman wandering around in dark places and lonely car parks alone-oblivious to the consequences. Great riddles that fell by the wayside. The promise of the knowledge therein contained by the original so-called criminal. I had no problem with the budget and enjoyed the suspense. I understood and can wax lyrical about the fool and found

Adrian Pauls role crucial and penetrating and then ????? Basically the story line and the script where good up to a point and that point was the last 10 minutes or so. What? Run out of ideas! Such a pity that this movie had to let us down so badly. It may not comprehend the meaning and I really did not expect the writers to understand it but I was hoping for an intellectual, if not spiritual ride and got a bump in the road"

בביקורת הזו יש מילים חיוביות בהקשר של הציפיה לסרט והתחלות מבטיחות בו, (Great and found Adrian Pauls role crucial and penetrating ,riddles וגם מילים שליליות המביעות אכזבה, וכנראה שלכן המודל טעה.

5.

*"Lowe returns to the nest after, yet another, failed relationship, to find he's been assigned to jury duty. It's in the plans to, somehow, get out of it, when he realizes the defendant is the girl he's had a serious crush on since the first grade.

Through living in the past by telling other people about his feelings towards this girl (played by Camp), Lowe remembers those feelings and does everything in his power to clear Camp of attempted murder, while staying away from the real bad guys at the same time, and succeeding in creating a successful film at the same time.

I've heard that St Augustine is the oldest city in the US, and I also know it has some ties to Ponce de Leon, so the backdrop is a good place to start. Unfortunately, it's the only thing good about this movie. The local police are inept, the judge is an idiot, and the defense counsel does everything in her power to make herself look like Joanie Cunningham! I don't know whether to blame the director for poor direction, or for just letting the cast put in such a hapless effort.

In short, this movie was so boring, I could not even sleep through it! 1 out of 10 stars!"*

בביקורת זה החצי הראשון הוא עם סנטימנט חיובי (או לפחות אי שלילי) ורק בחצי השני מופיעות מילים שליליות, ויכול להיות שזה מה שגרם למודל לטעות.

שאלה 6

צ"ל:

$$P(x_0 x_1 \dots x_n) = P(x_n) P(x_{n-1} | x_n) \dots P(x_0 | x_1) = P(x_0) P(x_1 | x_0) \dots P(x_n | x_{n-1})$$

לפי חוק בייס:

$$P(x_t | x_{t-1}) = \frac{P(x_{t-1} | x_t) \cdot P(x_t)}{P(x_{t-1})}$$

מכאן:

$$\begin{aligned} & P(x_0) P(x_1 | x_0) \dots P(x_n | x_{n-1}) \\ &= P(x_0) \frac{P(x_0 | x_1) \cdot P(x_1)}{P(x_0)} \cdot \frac{P(x_1 | x_2) \cdot P(x_2)}{P(x_1)} \cdot \dots \cdot \frac{P(x_{n-1} | x_n) \cdot P(x_n)}{P(x_{n-1})} \end{aligned}$$

נצמצם איברים ונקבל:

$$= P(x_0 | x_1) \cdot P(x_1 | x_2) \cdot \dots \cdot P(x_{n-1} | x_n) \cdot P(x_n)$$

כנדרש.