בס"ד

מגיש: מיכאל לוגסי
ת.ז: 305536575

דו"ח תרגיל בית 2

1. בוצע בקוד להלן פירוט פעולות הפונקציות

- **backward** - הפונקציה מבצבעת פיעפוע אחורה של ה loss לכלל שכבות הרשת, על מנת לבצע שיפור למשקולות הרשת בהתאם לקצב הלימוד, תוך התחשבות בחלק הטעות של כל משקולת בתוצאת ה loss.
- **loss_compute** – הפונקציה מבצעת חישוב של Cross-Entropy Loss, ע"י חיבור כל תוצאות המינוס של ה log של התוצאה (ההסתברות) של הקטגוריה שהתבקלה בחיזוי, וחלוקה שלו בכמות הדגימות. וכך למעשה היא מחשבת את ממוצע המרחק של הטעות מהאמת עבור כל תוצאות החוזי.
- **softmax** – הפונקציה מבצעת נירמול על פלט של הרשת, על מנת לייצג את תוצאת קלאסיפיקציה, בצורה מנורמלת ומייצגת סוג של הסתברות עבור כל קטגוריה.
הפונקציה בעצם מבצעת אקפוננט עבור כל אחת מהתוצאות של כל אחת מהקטגוריות ומחלקת בסכום האקספוננטים של תוצאות כל הקטגוריות.

2. להלן תוצאות אימון המודל עם וללא נירמול.
(ביצעת את האימונים על אותו המודל ללא שינוי מלבד הנדרש, לטובת ביצוע ההשוואה)

- עבור אימון **עם** נירמול, ניתן לראות שדיוק המודל עומד על כ 46%, גם לאחר משחק עם כמות השכבות והנוירונים, זה התוצאות פחות או יותר שקיבלתי.

```
---------------------------------------
Model Setup:
---------------------------------------
Data Set: ex2
Batch Size: 32
Normalize: True
Activation: relu
Epochs: 10
Learning Rate: 0.1
---------------------------------------
Model Results:
---------------------------------------
Model Acccurency: 46.61%
Train Acccurency: 45.16%
Train Loss: 1.4879457592331584
---------------------------------------
```

- עבור אימון **ללא** נירמול, ניתן לראות שדיוק המודל עומד על כ 10%, למרות שהוא בוצע על אותה הרשת בדיוק.
ניסיתי גם לשחק עם ה learning rate אבל לא מצאתי שיפור, אולי שינוי מבנה הרשת (כמות נוירונים ושכבות) אך זאת לא בדקתי.

```
---------------------------------------
Model Setup:
---------------------------------------
Data Set: ex2
Batch Size: 32
Normalize: False
Activation: relu
Epochs: 10
Learning Rate: 0.1
---------------------------------------
Model Results:
---------------------------------------
Model Acccurency: 10.0%
Train Acccurency: 6.45%
Train Loss: 2.310888567468281
---------------------------------------
```

3. לפני השיעור שהיה אתמול, הייתי אומר שאולי כדאי להוסיף עוד כמה שכבות טובות לרשת כדי שנוכל לקבל יותר עומק ואדפטיביות מהרשת למגוון רחב של תמונות או לעשות "נירמול" אחר לתמונה, כלומר עיבוד כלשהו, או אולי להשתמש בפונקציית אקטיבציה אחרת, שתטיב עם עבודה עם תמונות.
אחרי השיעור של אתמול, ברור לי שצריך לעשות קונבלוציה 😊

4. להלן תוצאות עבור ריצה עם הנתונים של תרגיל 1 (שוב על אותה רשת בדיוק, לטובת ביצוע ההשוואה, אך יתכן שעם התאמות בחלק מהפרמטים התוצאות היו משתפרות).
ניתן לראות שתוצאות די דומות למודל מהתרגיל הראשון (שם היה דיוק של כ 65% גם כן), אך כאן בוצעו משמעותית פחות איפוקים וגם ה learning rate הוא גם גדול הרבה יותר.

כמו כן, ניתן לראות שהמודל מצא "פתרון" טוב יותר (כמעט פי 2) מהבעיה של דאטה של תרגיל 2, למרות שהכמות דאטה שעליו אומן עבור הדאטה של התרגיל הראשון קטן משמעותית מהדאטה של התרגיל השני (אך כך גם כמות הפיצ'רים שלו קטנה משמעותית)

```
---------------------------------------
Model Setup:
---------------------------------------
Data Set: ex1
Batch Size: 32
Normalize: True
Activation: relu
Epochs: 10
Learning Rate: 0.1
---------------------------------------
Model Results:
---------------------------------------
Model Acccurency: 65.1%
Train Acccurency: 58.06%
Train Loss: 0.6958366671114682
---------------------------------------
```

5. להלן תוצאות הרשת לאחר החלפת פונקציית האקטיבציה לפונקציית sigmoid
(גם כאן בחרתי להשאיר את אותה רשת בדיוק, מלבד השינוי של פונקציית אקטיבציה, לטובת השוואה, ייתכן שעם שינוי חלק מהפרטמטרים התוצאות היו משתפרות)

- ניתן לראות את תוצאות המודל עבור הדאטה של תרגיל 2, ניתן לראות שאיכות המודל ירדה משמעותית ל 10% סה"כ, יתכן ששוב המודל הגיע למינומום מקומי, אך שוב משחק על ה learning rate לא עזר לי לראות שיפור.

```
---------------------------------------
Model Setup:
---------------------------------------
Data Set: ex2
Batch Size: 32
Normalize: True
Activation: sigmoid
Epochs: 10
Learning Rate: 0.1
---------------------------------------
Model Results:
---------------------------------------
Model Acccurency: 10.0%
Train Acccurency: 6.45%
Train Loss: 2.311052097333386
---------------------------------------
```

- ניתן לראות את תוצאות המודל עבור הדאטה של תרגיל 1, ניתן לראות שאיכות המודל נשארה פחות או יותר אותו הדבר בדיוק, כנראה שבמקרה של בעית הנתונים של תרגיל 1, אין כ"כ הבדל בין פונקציית האקטבציה של sigmoid או relu

```
---------------------------------------
Model Setup:
---------------------------------------
Data Set: ex1
Batch Size: 32
Normalize: True
Activation: sigmoid
Epochs: 10
Learning Rate: 0.1
---------------------------------------
Model Results:
---------------------------------------
Model Acccurency: 66.15%
Train Acccurency: 64.52%
Train Loss: 0.6553278162706836
---------------------------------------
```