

# Deep Learning Assignment 1

הקוד לכלל הפונקציות והאבלואציה כולל הבנוס בקובץ מחברת המצורף.

## דוח תוצאות:

עבור סעיף 4 (בלי batch normalization):  
התכנסות לאחר 88 איטרציות:  
פירוט התוצאות (תוצאות הדיוק בשורות האחרונות)

```
-----  
Iter: 1, epoch 1 cost 2.264, validation accuracy: 20.0%  
-----  
Iter: 1, epoch 2 cost 1.986, validation accuracy: 28.344%  
-----  
Iter: 1, epoch 3 cost 1.861, validation accuracy: 32.606%  
-----  
Iter: 1, epoch 4 cost 1.765, validation accuracy: 35.031%  
-----  
Iter: 1, epoch 5 cost 1.668, validation accuracy: 39.182%  
-----  
Iter: 1, epoch 6 cost 1.6, validation accuracy: 40.83%  
-----  
Iter: 1, epoch 7 cost 1.545, validation accuracy: 43.287%  
-----  
Iter: 1, epoch 8 cost 1.49, validation accuracy: 46.264%  
-----  
Iter: 1, epoch 9 cost 1.43, validation accuracy: 48.787%  
-----  
Iter: 1, epoch 10 cost 1.373, validation accuracy: 51.305%  
-----  
Iter: 2, epoch 1 cost 1.315, validation accuracy: 53.802%  
-----  
Iter: 2, epoch 2 cost 1.257, validation accuracy: 55.69%  
-----  
Iter: 2, epoch 3 cost 1.198, validation accuracy: 57.769%  
-----  
Iter: 2, epoch 4 cost 1.144, validation accuracy: 59.654%  
-----  
Iter: 2, epoch 5 cost 1.099, validation accuracy: 61.605%  
-----  
Iter: 2, epoch 6 cost 1.059, validation accuracy: 63.007%  
-----  
Iter: 2, epoch 7 cost 1.023, validation accuracy: 64.531%  
-----  
Iter: 2, epoch 8 cost 0.992, validation accuracy: 66.024%  
-----  
Iter: 2, epoch 9 cost 0.963, validation accuracy: 67.224%  
-----  
Iter: 2, epoch 10 cost 0.935, validation accuracy: 68.419%  
-----  
Iter: 3, epoch 1 cost 0.909, validation accuracy: 69.587%  
-----  
Iter: 3, epoch 2 cost 0.882, validation accuracy: 70.473%  
-----  
Iter: 3, epoch 3 cost 0.854, validation accuracy: 71.702%  
-----
```

Iter: 3, epoch 4 cost 0.825, validation accuracy: 73.448%  
-----  
Iter: 3, epoch 5 cost 0.797, validation accuracy: 74.398%  
-----  
Iter: 3, epoch 6 cost 0.773, validation accuracy: 75.651%  
-----  
Iter: 3, epoch 7 cost 0.753, validation accuracy: 76.53%  
-----  
Iter: 3, epoch 8 cost 0.736, validation accuracy: 77.214%  
-----  
Iter: 3, epoch 9 cost 0.72, validation accuracy: 77.743%  
-----  
Iter: 3, epoch 10 cost 0.706, validation accuracy: 78.449%  
-----  
Iter: 4, epoch 1 cost 0.693, validation accuracy: 78.975%  
-----  
Iter: 4, epoch 2 cost 0.68, validation accuracy: 79.276%  
-----  
Iter: 4, epoch 3 cost 0.669, validation accuracy: 79.733%  
-----  
Iter: 4, epoch 4 cost 0.659, validation accuracy: 80.148%  
-----  
Iter: 4, epoch 5 cost 0.648, validation accuracy: 80.475%  
-----  
Iter: 4, epoch 6 cost 0.638, validation accuracy: 80.674%  
-----  
Iter: 4, epoch 7 cost 0.629, validation accuracy: 81.038%  
-----  
Iter: 4, epoch 8 cost 0.62, validation accuracy: 81.267%  
-----  
Iter: 4, epoch 9 cost 0.612, validation accuracy: 81.554%  
-----  
Iter: 4, epoch 10 cost 0.605, validation accuracy: 81.701%  
-----  
Iter: 5, epoch 1 cost 0.597, validation accuracy: 82.038%  
-----  
Iter: 5, epoch 2 cost 0.59, validation accuracy: 82.185%  
-----  
Iter: 5, epoch 3 cost 0.583, validation accuracy: 82.402%  
-----  
Iter: 5, epoch 4 cost 0.576, validation accuracy: 82.551%  
-----  
Iter: 5, epoch 5 cost 0.57, validation accuracy: 82.739%  
-----  
Iter: 5, epoch 6 cost 0.563, validation accuracy: 82.769%  
-----  
Iter: 5, epoch 7 cost 0.558, validation accuracy: 83.138%  
-----  
Iter: 5, epoch 8 cost 0.552, validation accuracy: 83.229%  
-----  
Iter: 5, epoch 9 cost 0.546, validation accuracy: 83.423%  
-----  
Iter: 5, epoch 10 cost 0.541, validation accuracy: 83.512%  
-----  
Iter: 6, epoch 1 cost 0.535, validation accuracy: 83.592%  
-----  
Iter: 6, epoch 2 cost 0.531, validation accuracy: 83.94%  
-----  
Iter: 6, epoch 3 cost 0.525, validation accuracy: 84.087%  
-----  
Iter: 6, epoch 4 cost 0.52, validation accuracy: 84.249%  
-----  
Iter: 6, epoch 5 cost 0.515, validation accuracy: 84.346%

```
-----
Iter: 6, epoch 6 cost 0.51, validation accuracy: 84.44%
-----
Iter: 6, epoch 7 cost 0.505, validation accuracy: 84.617%
-----
Iter: 6, epoch 8 cost 0.501, validation accuracy: 84.817%
-----
Iter: 6, epoch 9 cost 0.496, validation accuracy: 84.917%
-----
Iter: 6, epoch 10 cost 0.491, validation accuracy: 84.999%
-----
Iter: 7, epoch 1 cost 0.487, validation accuracy: 85.061%
-----
Iter: 7, epoch 2 cost 0.482, validation accuracy: 85.357%
-----
Iter: 7, epoch 3 cost 0.478, validation accuracy: 85.419%
-----
Iter: 7, epoch 4 cost 0.474, validation accuracy: 85.486%
-----
Iter: 7, epoch 5 cost 0.47, validation accuracy: 85.674%
-----
Iter: 7, epoch 6 cost 0.466, validation accuracy: 85.7%
-----
Iter: 7, epoch 7 cost 0.462, validation accuracy: 85.842%
-----
Iter: 7, epoch 8 cost 0.458, validation accuracy: 85.933%
-----
Iter: 7, epoch 9 cost 0.454, validation accuracy: 86.04%
-----
Iter: 7, epoch 10 cost 0.451, validation accuracy: 86.14%
-----
Iter: 8, epoch 1 cost 0.446, validation accuracy: 86.169%
-----
Iter: 8, epoch 2 cost 0.443, validation accuracy: 86.344%
-----
Iter: 8, epoch 3 cost 0.44, validation accuracy: 86.418%
-----
Iter: 8, epoch 4 cost 0.437, validation accuracy: 86.44%
-----
Iter: 8, epoch 5 cost 0.433, validation accuracy: 86.464%
-----
Iter: 8, epoch 6 cost 0.43, validation accuracy: 86.632%
-----
Iter: 8, epoch 7 cost 0.427, validation accuracy: 86.819%
-----
Iter: 8, epoch 8 cost 0.424, validation accuracy: 86.823%
Results summary:
Batch size = 64
Train accuracy: 87.51%
Validation accuracy: 87.6%
Test accuracy: 87.4%
```

סעיף 5(עם batch normalization):  
התכנסות מהירה הרבה יותר (אחרי 43) איטרציות שהביאה גם לשיפור הדיוק:

```
-----
Iter: 1, epoch 1 cost 1.973, validation accuracy: 31.249%
-----
Iter: 1, epoch 2 cost 1.552, validation accuracy: 45.976%
-----
Iter: 1, epoch 3 cost 1.284, validation accuracy: 57.006%
-----
Iter: 1, epoch 4 cost 1.099, validation accuracy: 64.694%
```

-----  
Iter: 1, epoch 5 cost 0.968, validation accuracy: 71.957%  
-----  
Iter: 1, epoch 6 cost 0.85, validation accuracy: 78.11%  
-----  
Iter: 1, epoch 7 cost 0.744, validation accuracy: 81.748%  
-----  
Iter: 1, epoch 8 cost 0.649, validation accuracy: 84.449%  
-----  
Iter: 1, epoch 9 cost 0.575, validation accuracy: 86.283%  
-----  
Iter: 1, epoch 10 cost 0.517, validation accuracy: 87.57%  
-----  
Iter: 2, epoch 1 cost 0.477, validation accuracy: 88.676%  
-----  
Iter: 2, epoch 2 cost 0.447, validation accuracy: 89.451%  
-----  
Iter: 2, epoch 3 cost 0.418, validation accuracy: 90.021%  
-----  
Iter: 2, epoch 4 cost 0.397, validation accuracy: 90.48%  
-----  
Iter: 2, epoch 5 cost 0.381, validation accuracy: 90.802%  
-----  
Iter: 2, epoch 6 cost 0.364, validation accuracy: 91.045%  
-----  
Iter: 2, epoch 7 cost 0.351, validation accuracy: 91.326%  
-----  
Iter: 2, epoch 8 cost 0.343, validation accuracy: 91.686%  
-----  
Iter: 2, epoch 9 cost 0.332, validation accuracy: 91.873%  
-----  
Iter: 2, epoch 10 cost 0.327, validation accuracy: 91.982%  
-----  
Iter: 3, epoch 1 cost 0.319, validation accuracy: 92.101%  
-----  
Iter: 3, epoch 2 cost 0.314, validation accuracy: 92.243%  
-----  
Iter: 3, epoch 3 cost 0.307, validation accuracy: 92.392%  
-----  
Iter: 3, epoch 4 cost 0.301, validation accuracy: 92.45%  
-----  
Iter: 3, epoch 5 cost 0.294, validation accuracy: 92.579%  
-----  
Iter: 3, epoch 6 cost 0.295, validation accuracy: 92.652%  
-----  
Iter: 3, epoch 7 cost 0.291, validation accuracy: 92.785%  
-----  
Iter: 3, epoch 8 cost 0.28, validation accuracy: 92.856%  
-----  
Iter: 3, epoch 9 cost 0.281, validation accuracy: 92.902%  
-----  
Iter: 3, epoch 10 cost 0.276, validation accuracy: 92.96%  
-----  
Iter: 4, epoch 1 cost 0.274, validation accuracy: 93.03%  
-----  
Iter: 4, epoch 2 cost 0.272, validation accuracy: 93.083%  
-----  
Iter: 4, epoch 3 cost 0.269, validation accuracy: 93.082%

**Results summary:**

**Batch size = 64**

**Train accuracy: 94.04%**

**Validation accuracy: 93.908%**

סעיף 6(הוספת L2 regularization ובלי batch normalization):  
**הסבר שינויים בקוד בבונס** – כדי לתמוך בגולריזציה אנחנו צריכים להוסיף את הרגולריזציה לפונקציית ההפסד שלנו ובעקבות כך גם לעדכן את המשקולות בהתאם.  
כפי שלמדנו בכיתה, ספציפית עבור רגולריזציה 2 – נוסיף לפונקציית ההפסד את סכום המשקולות בריבוע כפול מקדם הרגולריזציה חלקי 2. עבור עדכון המשקולות – מגזירת החלק של הרגולריזציה בפונקציית ההפסד נקבל שצריך להוסיף את מקדם הרגולריזציה כפול ערך המשקולות (בלי העלאה בריבוע)  
ניתוח תוצאות (עבור מקדם רגולריזציה = 1) – אפשר לראות את התוצאות האלה גם במחברת שהוספנו, נסתכל על תוצאות השכבה הראשונה.  
התוצאות של המשקולות שקיבלנו עבור הרצה עם רגולריזציה (כמו בסעיף 4 – ללא נרמול):

```
✓ [59] 1 parameters['W'][0]  
0s  
array([[ 0.24835708, -0.06913215,  0.32384427, ..., -0.66767218,  
         0.19009893,  0.30529287],  
       [ 0.27989522,  0.54039036,  0.41696108, ..., -0.14066378,  
         0.03349536,  0.25796961],  
       [-0.78127293, -0.26452634,  0.39713234, ..., -0.19600632,  
         0.52996819,  0.30850297],  
       ...,  
       [ 0.02769846,  0.21752419, -0.41420161, ..., -1.11296771,  
        -0.39377572, -0.23497894],  
       [-0.11799005, -0.24497201, -1.27501315, ..., -0.81072631,  
        -0.3751294 ,  0.2941184 ],  
       [ 0.36096063, -0.32585207, -0.06707666, ..., -0.47945705,  
        -0.36671258, -0.00550316]])
```

התוצאות של המשקולות ללא רגולריזציה עבור השכבה האחרונה:

```
1 parameters_12['W'][0]

array([[ 0.00103693, -0.00202308,  0.00057438, ..., -0.01789345,
         0.01167333,  0.01206085],
       [-0.01791017,  0.00131053,  0.01696928, ...,  0.02101566,
         0.00231083, -0.00573269],
       [-0.02229284,  0.00412522,  0.00324807, ...,  0.02210704,
        -0.01299997,  0.00461194],
       ...,
       [ 0.00267106, -0.00673723, -0.00825559, ...,  0.00824123,
         0.0119845 ,  0.01926042],
       [ 0.00041538, -0.00593797, -0.01522488, ...,  0.01309775,
        -0.00405748,  0.02241561],
       [-0.00826079, -0.00404636,  0.02323034, ...,  0.00429275,
         0.00746295, -0.02402202]])
```

ניתן לראות שערכי המשקולות קרובות הרבה יותר לאפס כמו שהיינו מצפים, בנוסף תוצאות הדיוק של הרשת הזאת גבוהות יותר מאשר כשלא השתמשנו ברגולריזציה וכנ"ל גם הדיוק שלה.

תוצאות התכנסות ודיוק המודל עם רגולריזציה (ללא נרמול):

-----  
Iter: 1, epoch 1 cost 30.092, validation accuracy: 11.792%

-----  
Iter: 1, epoch 2 cost 24.735, validation accuracy: 15.188%

-----  
Iter: 1, epoch 3 cost 20.373, validation accuracy: 19.43%

-----  
Iter: 1, epoch 4 cost 16.779, validation accuracy: 25.288%

-----  
Iter: 1, epoch 5 cost 13.821, validation accuracy: 33.686%

-----  
Iter: 1, epoch 6 cost 11.235, validation accuracy: 46.258%

-----  
Iter: 1, epoch 7 cost 9.019, validation accuracy: 60.831%

-----  
Iter: 1, epoch 8 cost 7.35, validation accuracy: 70.252%

-----  
Iter: 1, epoch 9 cost 6.043, validation accuracy: 76.377%

-----  
Iter: 1, epoch 10 cost 5.005, validation accuracy: 79.931%

-----  
Iter: 2, epoch 1 cost 4.175, validation accuracy: 82.4%

-----  
Iter: 2, epoch 2 cost 3.509, validation accuracy: 84.1%

-----  
Iter: 2, epoch 3 cost 2.972, validation accuracy: 85.392%

-----  
Iter: 2, epoch 4 cost 2.539, validation accuracy: 86.138%

-----  
Iter: 2, epoch 5 cost 2.189, validation accuracy: 86.86%

-----  
Iter: 2, epoch 6 cost 1.907, validation accuracy: 87.504%

-----  
Iter: 2, epoch 7 cost 1.678, validation accuracy: 88.082%

-----  
Iter: 2, epoch 8 cost 1.492, validation accuracy: 88.538%  
-----  
Iter: 2, epoch 9 cost 1.34, validation accuracy: 89.056%  
-----  
Iter: 2, epoch 10 cost 1.218, validation accuracy: 89.354%  
-----  
Iter: 3, epoch 1 cost 1.118, validation accuracy: 89.79%  
-----  
Iter: 3, epoch 2 cost 1.038, validation accuracy: 89.912%  
-----  
Iter: 3, epoch 3 cost 0.972, validation accuracy: 90.288%  
-----  
Iter: 3, epoch 4 cost 0.92, validation accuracy: 90.431%  
-----  
Iter: 3, epoch 5 cost 0.876, validation accuracy: 90.643%  
-----  
Iter: 3, epoch 6 cost 0.842, validation accuracy: 90.858%  
-----  
Iter: 3, epoch 7 cost 0.814, validation accuracy: 90.971%  
-----  
Iter: 3, epoch 8 cost 0.79, validation accuracy: 91.108%  
-----  
Iter: 3, epoch 9 cost 0.771, validation accuracy: 91.265%  
-----  
Iter: 3, epoch 10 cost 0.756, validation accuracy: 91.38%  
-----  
Iter: 4, epoch 1 cost 0.743, validation accuracy: 91.486%  
-----  
Iter: 4, epoch 2 cost 0.732, validation accuracy: 91.589%  
-----  
Iter: 4, epoch 3 cost 0.724, validation accuracy: 91.639%  
-----  
Iter: 4, epoch 4 cost 0.716, validation accuracy: 91.692%  
-----  
Iter: 4, epoch 5 cost 0.71, validation accuracy: 91.836%  
-----  
Iter: 4, epoch 6 cost 0.705, validation accuracy: 91.736%  
**Results summary:**  
**Batch size = 64**  
**Train accuracy: 92.294%**  
**Validation accuracy: 91.958%**  
**Test accuracy: 91.96%**