به نام خدا

گزارش پروژه اول درس یادگیری عمیق

دکتر سید ابوالقاسم میرروشندل

تاریخ تحویل: ۱۴۰۱/۰۱/۱۵

نام و نام خانوادگی	فاطمه کمانی
آدرس ایمیل	anitakamani@outlook.com
شماره دانشجویی	970122680033

- **نتایج به‌ازای مقداردهی اولیه رندوم**

Random initialization:

k = 1000
alpha = 0.1

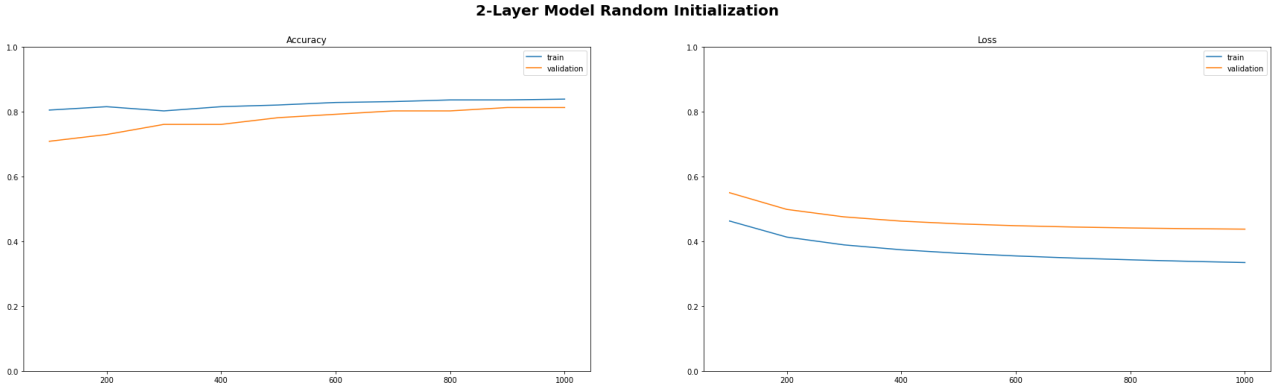
Number of Layers	Accuracy and Loss of train set according to Min train Loss	Accuracy and Loss of val set according to Min val Loss	Accuracy and Loss of train set according to Max train Accuracy	Accuracy and Loss of train set according to Max val Accuracy
2	Accuracy: 0.8385416666666666 Loss: 0.334180963296714	Accuracy: 0.8125 Loss: 0.43724306517513023	Accuracy: 0.8385416666666666 Loss: 0.3416983605202647	Accuracy: 0.8125 Loss: 0.43918698146776114
3	Accuracy: 0.8645833333333334 Loss: 0.2960751313068038	Accuracy: 0.8020833333333334 Loss: 0.4312179848829363	Accuracy: 0.8645833333333334 Loss: 0.2962985847697907	Accuracy: 0.8125 Loss: 0.43268440375523975
5	Accuracy: 0.7317708333333334 Loss: 0.5814894426274395	Accuracy: 0.59375 Loss: 0.6754653742870771	Accuracy: 0.7317708333333334 Loss: 0.688163692834787	Accuracy: 0.59375 Loss: 0.691166596804643

- **تعداد لایه=۲:**
  - **شکل خروجی کد (Loss و Accuracy) مجموعه آموزش**
  - **شکل خروجی کد (Loss و Accuracy) مجموعه ارزیابی**

2-Layer Model Random Initialization

```
[798]  1  accuracy_loss_random_2,report_k_vals = model(X_train, y_train, X_val, y_val, architecture=architecture_2, alpha=alpha, k=k, report_k=report_k, random=True)
       2  draw_pot("2-Layer Model Random Initialization", report_k_vals)

100%|██████████| 1000/1000 [00:00<00:00, 3748.82it/s]


itrations  :  [ 100  200  300  400  500  600  700  800  900 1000]


loss_train  :  [0.46235197 0.41273914 0.38866299 0.3735793  0.3628334  0.35461041
 0.34803574 0.34262239 0.33807035 0.33418096]


accuracy_train  :  [0.8046875  0.81510417 0.80208333 0.81510417 0.8203125  0.828125
 0.83072917 0.8359375  0.8359375  0.83854167]


loss_val  :  [0.54966268 0.49817623 0.47509826 0.46198602 0.45359966 0.4478814
 0.44383817 0.44092038 0.43879416 0.43724307]


accuracy_val  :  [0.70833333 0.72916667 0.76041667 0.76041667 0.78125    0.79166667
 0.80208333 0.80208333 0.8125     0.8125     ]
```

- **نمودار تغییر Loss مجموعه آموزش**
- **نمودار تغییر Loss مجموعه ارزیابی**
- **نمودار تغییر Accuracy مجموعه آموزش**
- **نمودار تغییر Accuracy مجموعه ارزیابی**



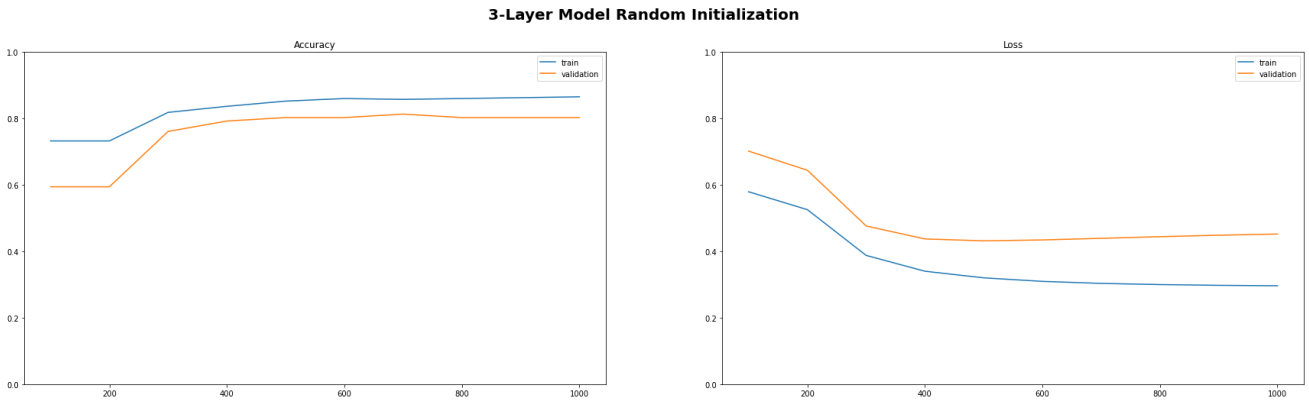
- تعداد لایه=۳:
  - شکل خروجی کد (Loss و Accuracy) مجموعه آموزش
  - شکل خروجی کد (Loss و Accuracy) مجموعه ارزیابی

▾ 3-Layer Model Random Initialization

```
1  accuracy_loss_random_3, report_k_vals = model(X_train, y_train, X_val, y_val, architecture=architecture_3, alpha=alpha, k=k, report_k=report_k, random=True)
2  draw_pot("3-Layer Model Random Initialization", report_k_vals)
```

```
100%|██████████| 1000/1000 [00:00<00:00, 1285.84it/s]

itrations  :  [ 100  200  300  400  500  600  700  800  900 1000]

loss_train  :  [0.57883254 0.52497274 0.38739406 0.33978241 0.31989889 0.30922276
 0.30315634 0.29959075 0.29742651 0.29607513]

accuracy_train  :  [0.73177083 0.73177083 0.81770833 0.8359375  0.8515625  0.859375
 0.85677083 0.859375   0.86197917 0.86458333]

loss_val  :  [0.70110272 0.64356769 0.47594378 0.43691515 0.43121871 0.4339114
 0.4387489  0.44371737 0.44808191 0.45170119]

accuracy_val  :  [0.59375    0.59375    0.76041667 0.79166667 0.80208333 0.80208333
 0.8125     0.80208333 0.80208333 0.80208333]
```

  - نمودار تغییر Loss مجموعه آموزش
  - نمودار تغییر Loss مجموعه ارزیابی
  - نمودار تغییر Accuracy مجموعه آموزش
  - نمودار تغییر Accuracy مجموعه ارزیابی



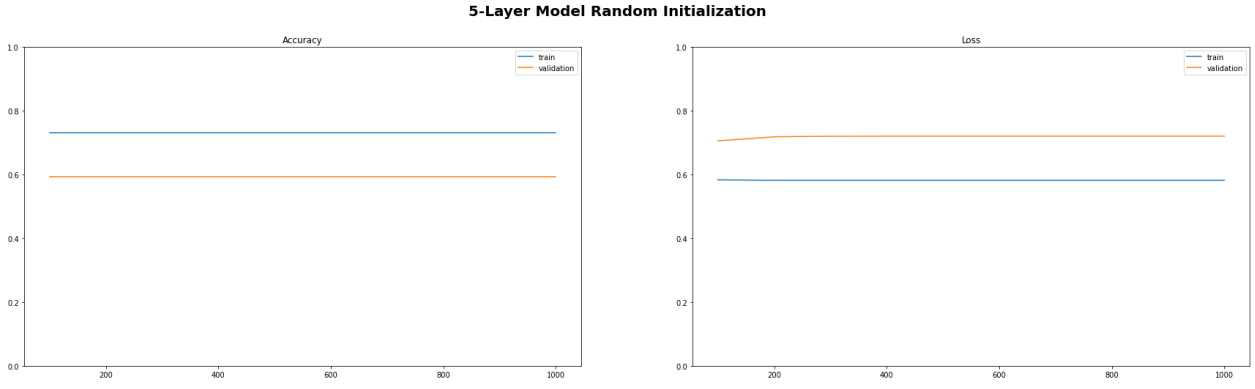
- تعداد لایه=۵:
  - شکل خروجی کد (Loss و Accuracy) مجموعه آموزش
  - شکل خروجی کد (Loss و Accuracy) مجموعه ارزیابی

5-Layer Model Random Initialization

```
accuracy_loss_random_5,report_k_vals = model(X_train, y_train, X_val, y_val, architecture=architecture_5, alpha=alpha, k=k, report_k=report_k, random=True)
draw_pot("5-Layer Model Random Initialization", report_k_vals)
```

```
100%|██████████| 1000/1000 [00:01<00:00, 638.73it/s]

itrations  :  [ 100  200  300  400  500  600  700  800  900 1000]

loss_train  :  [0.58291426 0.5815148  0.58148994 0.58148947 0.58148946 0.58148945
 0.58148945 0.58148945 0.58148945 0.58148944]

accuracy_train  :  [0.73177083 0.73177083 0.73177083 0.73177083 0.73177083 0.73177083
 0.73177083 0.73177083 0.73177083 0.73177083]

loss_val  :  [0.70522997 0.71786398 0.71971335 0.71996978 0.7200051  0.72000996
 0.72001062 0.72001071 0.72001072 0.72001072]

accuracy_val  :  [0.59375 0.59375 0.59375 0.59375 0.59375 0.59375 0.59375 0.59375 0.59375 0.59375
 0.59375]
```

- نمودار تغییر Loss مجموعه آموزش
- نمودار تغییر Loss مجموعه ارزیابی
- نمودار تغییر Accuracy مجموعه آموزش
- نمودار تغییر Accuracy مجموعه ارزیابی



- **نتایج به‌ازای مقداردهی اولیه صفر**

Zero initialization:

k = 1000
alpha = 0.1

Number of Layers	Accuracy and Loss of train set according to Min train Loss	Accuracy and Loss of val set according to Min val Loss	Accuracy and Loss of train set according to Max train Accuracy	Accuracy and Loss of train set according to Max val Accuracy
2	Accuracy: 0.8385416666666666 Loss: 0.3341298061400587	Accuracy: 0.8125 Loss: 0.4372580017286208	Accuracy: 0.8385416666666666 Loss: 0.34183055533922524	Accuracy: 0.8125 Loss: 0.4390021899421154
3	Accuracy: 0.7317708333333334 Loss: 0.5814895633418021	Accuracy: 0.59375 Loss: 0.6754659858598195	Accuracy: 0.7317708333333334 Loss: 0.6878425542876629	Accuracy: 0.59375 Loss: 0.6910414746435658
5	Accuracy: 0.7317708333333334 Loss: 0.5814895633418021	Accuracy: 0.59375 Loss: 0.6754659858598195	Accuracy: 0.7317708333333334 Loss: 0.6878425542876629	Accuracy: 0.59375 Loss: 0.6910414746435658

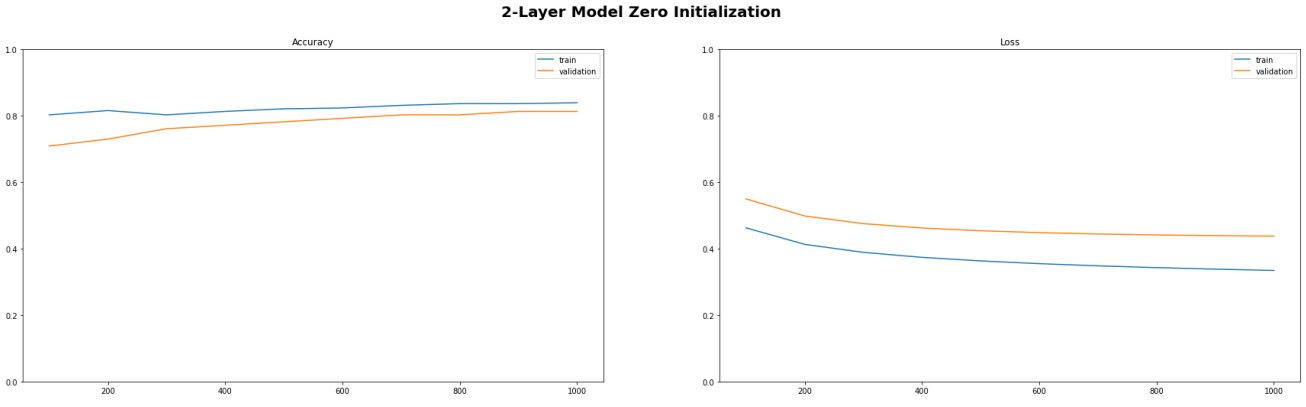
- تعداد لایه=۲:
  - شکل خروجی کد (Loss و Accuracy) مجموعه آموزش
  - شکل خروجی کد (Loss و Accuracy) مجموعه ارزیابی

2-Layer Model Zero Initialization

```
accuracy_loss_zero_2, report_k_vals= model(X_train, y_train, X_val, y_val, architecture=architecture_2, alpha=alpha, k=k, report_k=report_k, random=False)
draw_pot("2-Layer Model Zero Initialization", report_k_vals)
```

```
100%|██████████| 1000/1000 [00:00<00:00, 3484.24it/s]

itrations  :  [ 100  200  300  400  500  600  700  800  900 1000]

loss_train  :  [0.46222579 0.41257588 0.38852168 0.37346388 0.36273872 0.35453095
 0.34796716 0.34256158 0.33801512 0.33412981]

accuracy_train  :  [0.80208333 0.81510417 0.80208333 0.8125     0.8203125  0.82291667
 0.83072917 0.8359375  0.8359375  0.83854167]

loss_val  :  [0.54928815 0.49785582 0.47487404 0.46183975 0.45350957 0.44783043
 0.44381408 0.44091462 0.4388008  0.437258  ]

accuracy_val  :  [0.70833333 0.72916667 0.76041667 0.77083333 0.78125    0.79166667
 0.80208333 0.80208333 0.8125     0.8125    ]
```

- نمودار تغییر Loss مجموعه آموزش
- نمودار تغییر Loss مجموعه ارزیابی
- نمودار تغییر Accuracy مجموعه آموزش
- نمودار تغییر Accuracy مجموعه ارزیابی



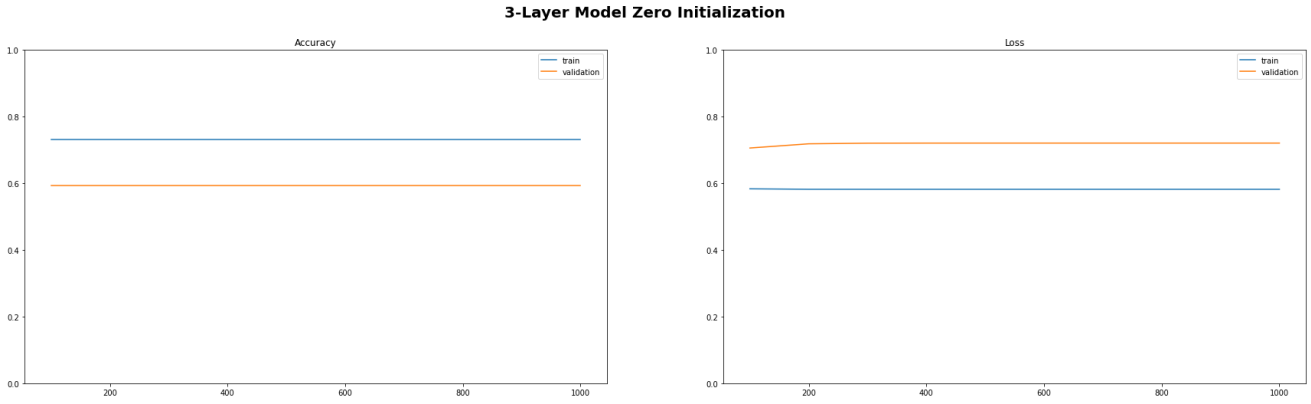
- تعداد لایه=۳:
  - شکل خروجی کد (Loss و Accuracy) مجموعه آموزش
  - شکل خروجی کد (Loss و Accuracy) مجموعه ارزیابی

3-Layer Model Zero Initialization

```
1  accuracy_loss_zero_3, report_k_vals = model(X_train, y_train, X_val, y_val, architecture=architecture_3, alpha=alpha, k=k, report_k=report_k, random=False)
2  draw_pot("3-Layer Model Zero Initialization", report_k_vals)
```

```
100%|██████████| 1000/1000 [00:00<00:00, 1371.92it/s]

itrations  :  [ 100  200  300  400  500  600  700  800  900 1000]

loss_train  :  [0.58291262 0.58151491 0.58149004 0.58148957 0.58148956 0.58148956
 0.58148956 0.58148956 0.58148956 0.58148956]

accuracy_train  :  [0.73177083 0.73177083 0.73177083 0.73177083 0.73177083 0.73177083
 0.73177083 0.73177083 0.73177083 0.73177083]

loss_val  :  [0.70523797 0.71786345 0.71971308 0.71996977 0.72000516 0.72001004
 0.72001071 0.7200108  0.72001081 0.72001082]

accuracy_val  :  [0.59375 0.59375 0.59375 0.59375 0.59375 0.59375 0.59375 0.59375 0.59375
 0.59375]
```

- نمودار تغییر Loss مجموعه آموزش
- نمودار تغییر Loss مجموعه ارزیابی
- نمودار تغییر Accuracy مجموعه آموزش
- نمودار تغییر Accuracy مجموعه ارزیابی



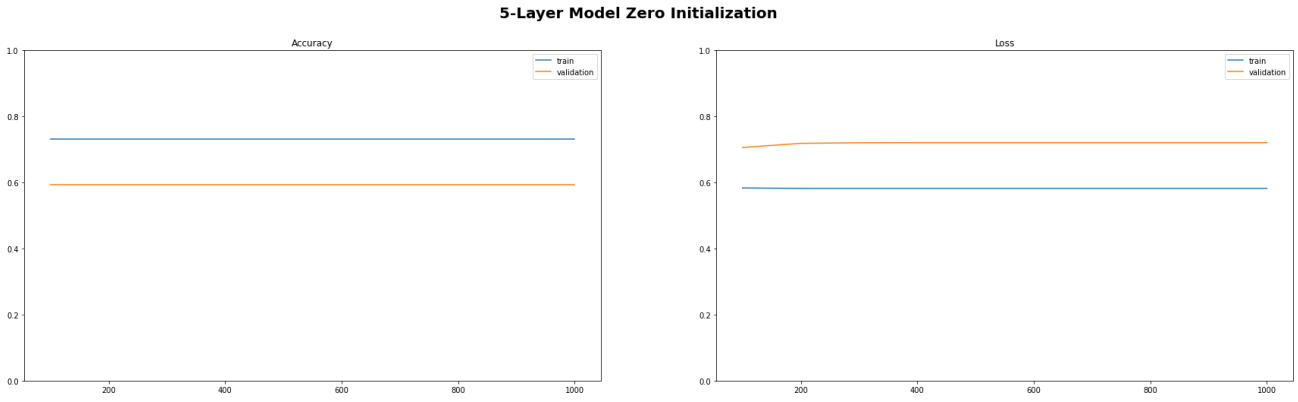
- تعداد لایه=۵:
  - شکل خروجی کد (Loss و Accuracy) مجموعه آموزش
  - شکل خروجی کد (Loss و Accuracy) مجموعه ارزیابی

5-Layer Model Zero Initialization

```
1  accuracy_loss_zero_5, report_k_vals= model(X_train, y_train, X_val, y_val, architecture=architecture_5, alpha=alpha, k=k, report_k=report_k, random=False)
2  draw_pot("5-Layer Model Zero Initialization", report_k_vals)
```

```
100%|██████████| 1000/1000 [00:01<00:00, 674.49it/s]

itrations  :  [ 100  200  300  400  500  600  700  800  900 1000]

loss_train  :  [0.58291262 0.58151491 0.58149004 0.58148957 0.58148956 0.58148956
 0.58148956 0.58148956 0.58148956 0.58148956]

accuracy_train  :  [0.73177083 0.73177083 0.73177083 0.73177083 0.73177083 0.73177083
 0.73177083 0.73177083 0.73177083 0.73177083]

loss_val  :  [0.70523797 0.71786345 0.71971308 0.71996977 0.72000516 0.72001004
 0.72001071 0.7200108  0.72001081 0.72001082]

accuracy_val  :  [0.59375 0.59375 0.59375 0.59375 0.59375 0.59375 0.59375 0.59375 0.59375
 0.59375]
```

  - نمودار تغییر Loss مجموعه آموزش
  - نمودار تغییر Loss مجموعه ارزیابی
  - نمودار تغییر Accuracy مجموعه آموزش
  - نمودار تغییر Accuracy مجموعه ارزیابی



- توضیحات تکمیلی:-

- **نتایج بهبود مدل (نمره مثبت):**

- Xavier initialization:

  we know that z is equal to:

$$Z = b + w1 \cdot x1 + w2 \cdot x2 + \ldots + wn \cdot xn$$

  while using Relu function we can set the variance to:

$$\sigma^2 = \frac{1}{size(l)}$$

  and then multiply this value in the weight matrix

$$W_l = np.random.randn(shape) * \sqrt{\frac{2}{size(l-1)}}$$

با ضرب این مقدار در ماتریس وزن‌ها، در شبکه‌های بزرگ‌تر شاهد بهبود نتایج هستیم:

در شبکه‌ی ۲ لایه‌ی زیر شاهد بدتر شدن Accuracy و Loss هستیم:

```
Random 0.01 vs xavier initialization of 2-layer NN:

Random initialization:
```

Number of Layers	Train Loss	Train Accuracy	Val Loss	Val Accuracy
0.01 Constant	0.334181	0.838542	0.437243	0.8125
Xavier Initializatio	0.33677	0.84375	0.438222	0.791667

```
Loss was reduced by  -0.22 %

Accuracy was improved by  -2.56 %
```

در شبکه ی ۳ لایه ی زیر شاهد بهبود هستیم اما مقدار آن چشمگیر نیست:

```
Random 0.01 vs xavier initialization of 3-layer NN:

Random initialization:
```

Number of Layers	Train Loss	Train Accuracy	Val Loss	Val Accuracy
0.01 Constant	0.296075	0.864583	0.451701	0.802083
Xavier Initializatio	0.259952	0.877604	0.422054	0.833333

```
Loss was reduced by  6.56 %

Accuracy was improved by  3.9 %
```

اما زمانی که به شبکه‌ی ۵ لایه می‌رسیم شاهد بهبود چشمگیر Accuracy و Loss هستیم.

```
Random 0.01 vs xavier initialization of 5-layer NN:
```

Number of Layers	Train Loss	Train Accuracy	Val Loss	Val Accuracy
0.01 Constant	0.581489	0.731771	0.720011	0.59375
Xavier Initializatio	0.149958	0.929688	0.37511	0.875

```
Loss was reduced by  47.9 %

Accuracy was improved by  47.37 %
```

این بهبود چشمگیر در شبکه‌های بزرگ‌تر باعث می‌شود تا ما نسبت به استفاده از این متد مصمم باشیم.