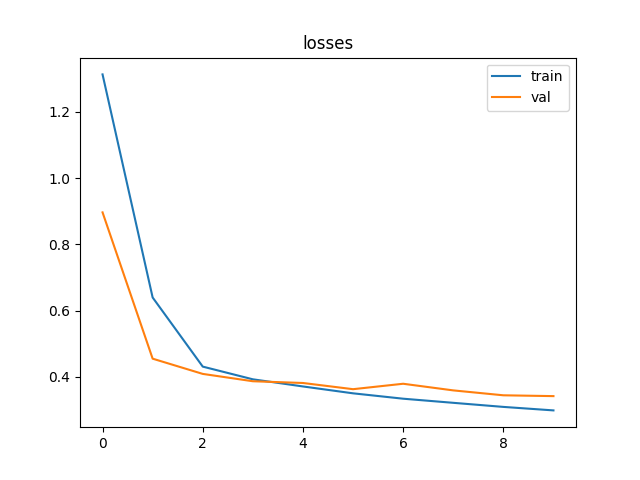
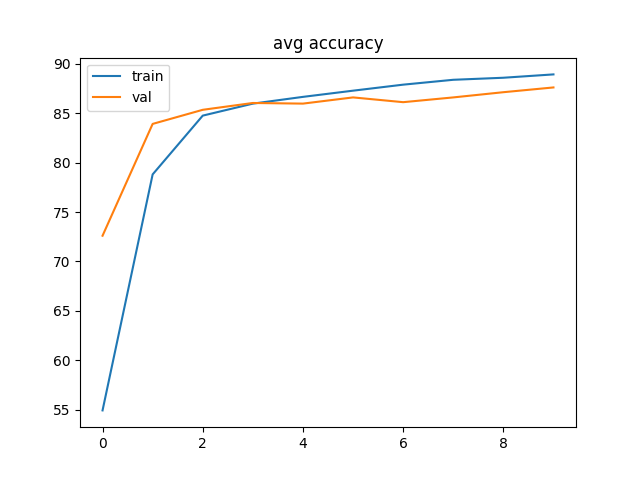
**דו"ח תרגיל 4**

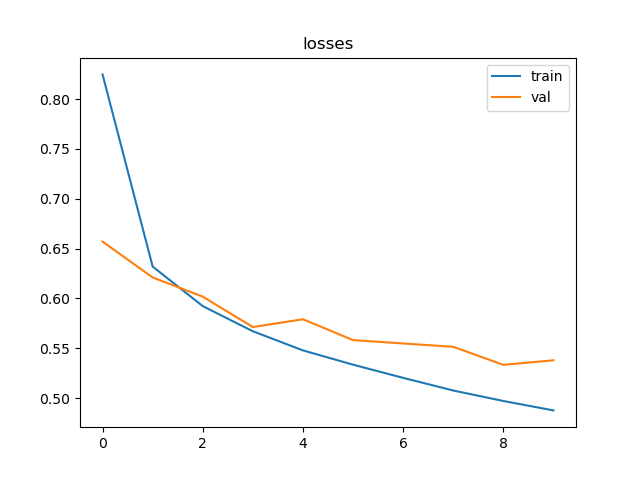
**מודל A**

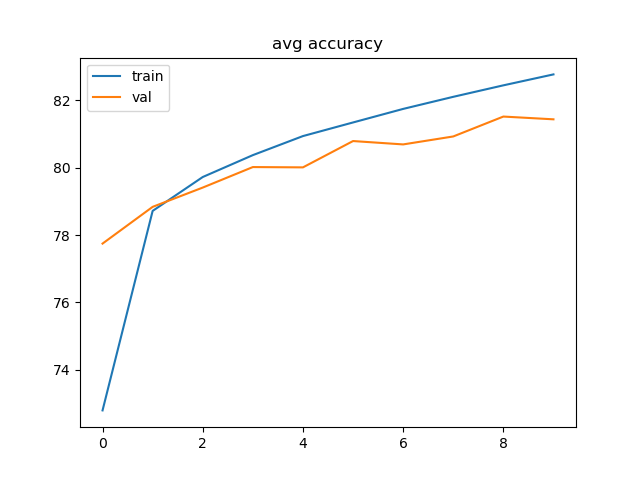


1. 
2. דיוק על test set של Fashion Mnist- 86.34%
3. היפר פרמטרים:

* Learning rate – 0.01
* Momentum (sgd) – 0.9
* Batch size – 64

**מודל B:**

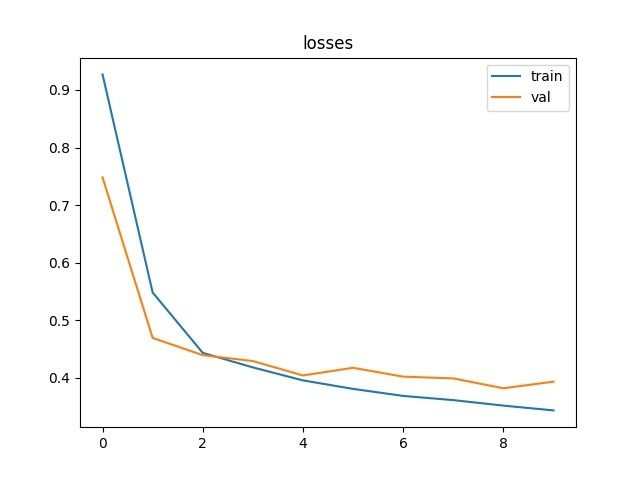


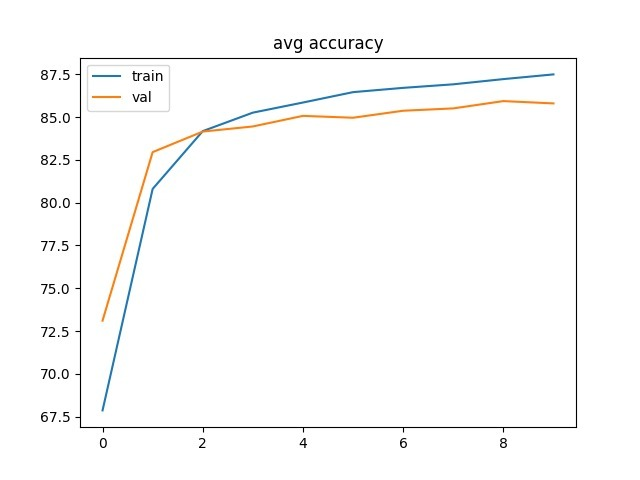


1. דיוק על test set של Fashion Mnist- 80.86%
2. היפר פרמטרים:

* Learning rate – 0.001
* Batch size – 64

**מודל C:**

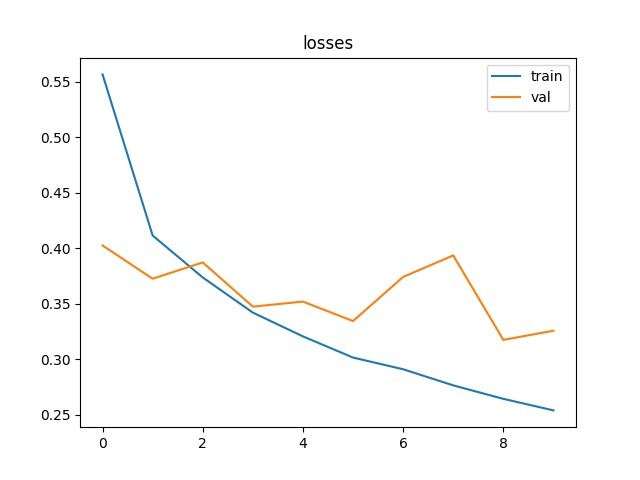
1.

2.

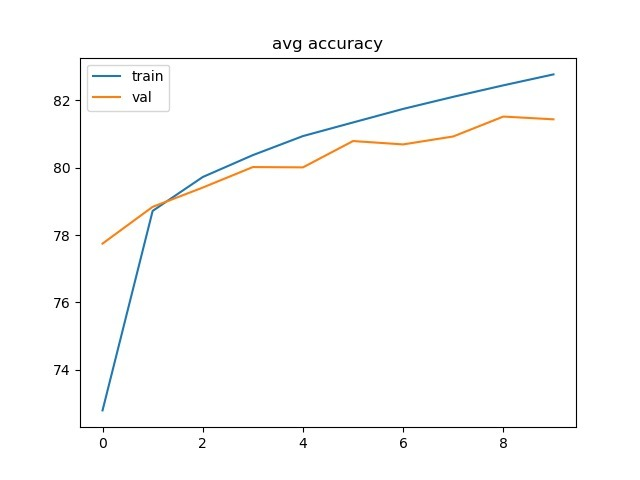
1. דיוק על test set של Fashion Mnist- 85%
2. היפר פרמטרים:

* Learning rate – 0.001
* Dropout – [0.3,0.1]
* Batch size – 64

**מודל D1 (batch norm מופעל לפני פונקציית האקטיבציה):**



1.



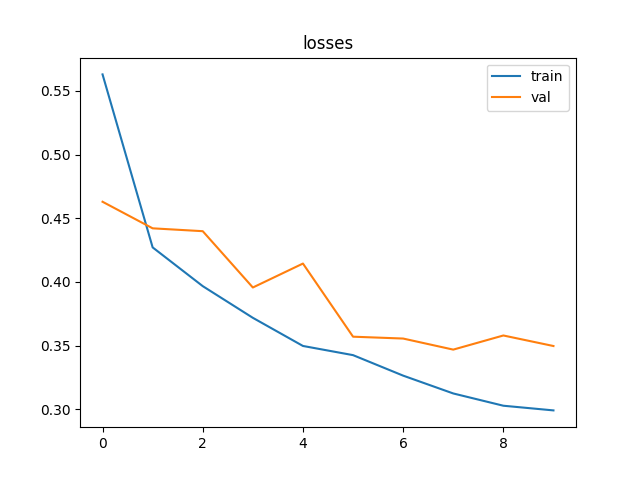
2.

1. דיוק על test set של Fashion Mnist- 87.5
2. היפר פרמטרים:

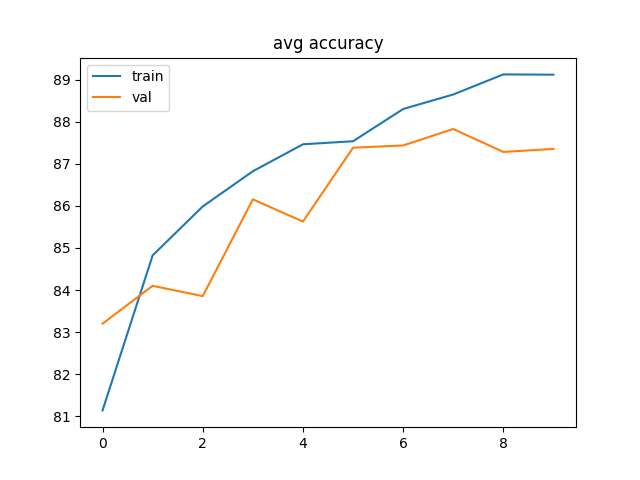
* Learning rate – 0.01
* Batch size – 64

**מודל D2 (batch norm מופעל אחרי פונקציית האקטיבציה):**

לא ראינו הבדל משמעותי בין הביצועים של batchnorm לפני האקטיבציה לבין הביצועים אחרי.

 1.

2.

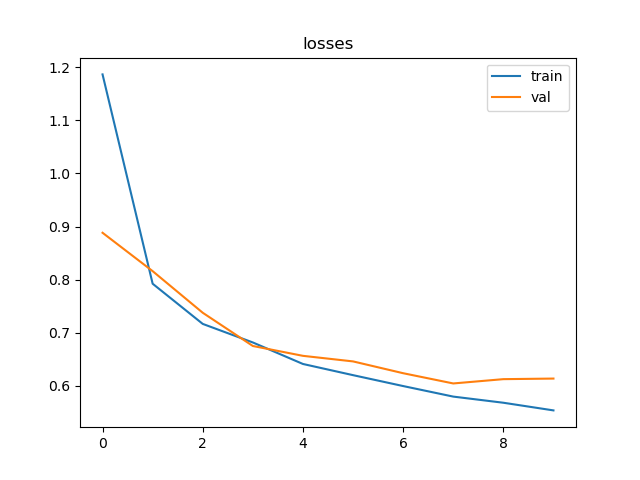


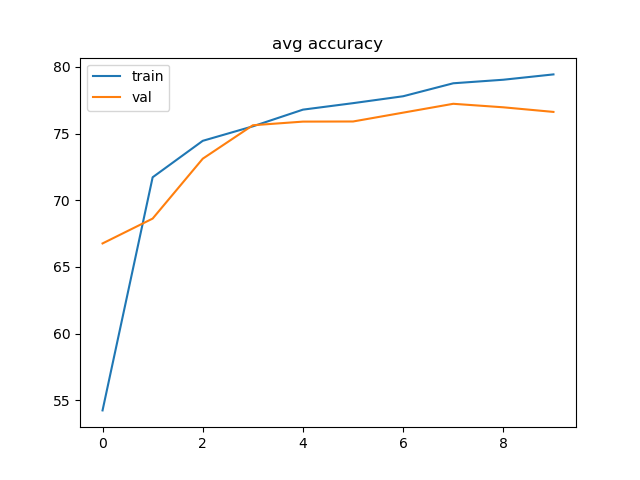
1. דיוק על test set של Fashion Mnist- 86.82%
2. היפר פרמטרים:

* Learning rate – 0.01
* Batch size – 64

**מודל E:**

1.

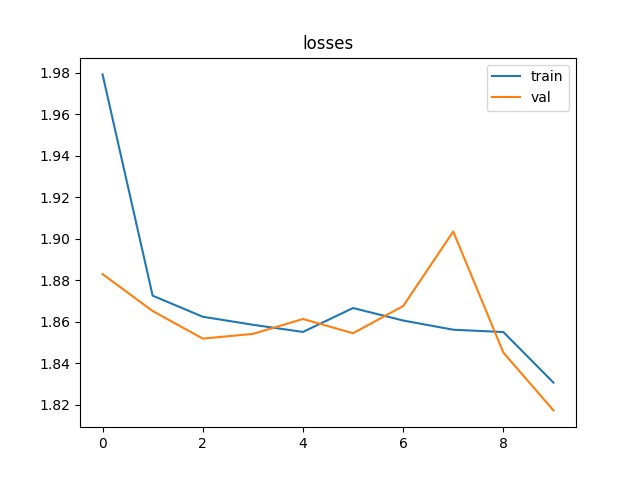


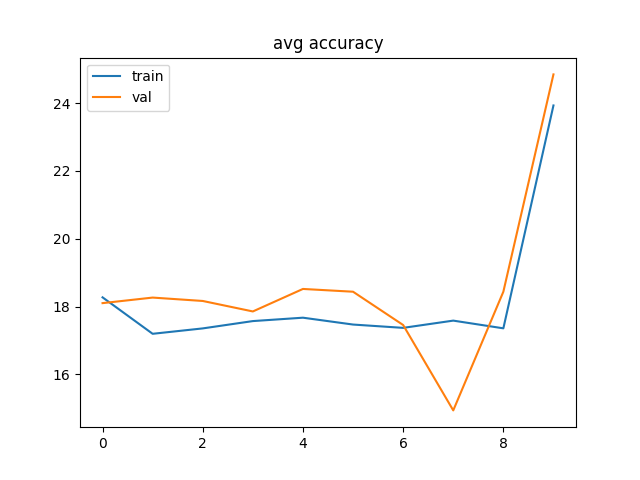


1. דיוק על test set של Fashion Mnist- 76.61%
2. היפר פרמטרים:

* Learning rate – 0.001
* Batch size – 64

**מודל F:**

1 .



2.

1. דיוק על test set של Fashion Mnist- 24.18%
2. היפר פרמטרים:

* Learning rate – 0.01
* Batch size – 64

**הארכיטקטורה של המודל הטוב ביותר:**

יצרנו 3 שכבות fully-conected בגדלים: 1024, 1024, 256

השתמשנו בbatchnorm בין כל שכבה, והשתמשנו בrelu כפונקציית אקטיבציה.

את השכבה האחרונה העברנו בsoftmax

היפר פרמטרים שהשתמשנו בהם:

Learning rate-0.001 (באמצע האימון הורדנו את הLr ל-1e-5)

Batch size – 64

Epochs - 20

השתמשנו בAdam כאופטימייזר

גרפים של ה-loss וה-accuracy:

Chart, line chart

Description automatically generatedChart, line chart

Description automatically generated