

Batch Size & Weight Initialization



[Link del Colaboratory](#)

Nil Torrents
Roger Vallejo
Eugeni Valentí
Adrián Rodríguez

Índex

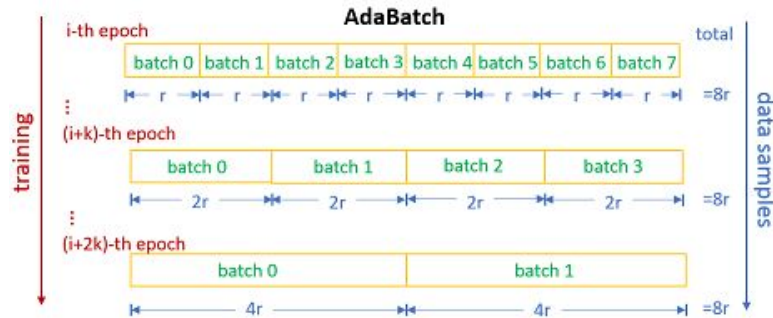


1. Conceptes bàsics
2. MNIST - Batch Size
3. MNIST - Weight Initialization
4. CIFAR10 B&W
5. Conclusions

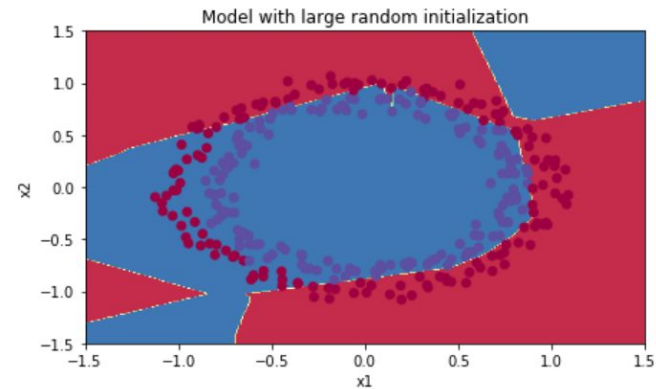


Conceptes Bàsics

Batch Size



Weight Initialization



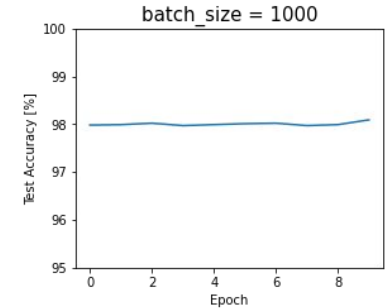
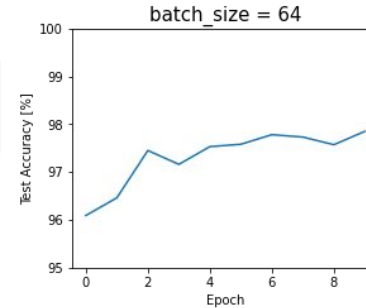
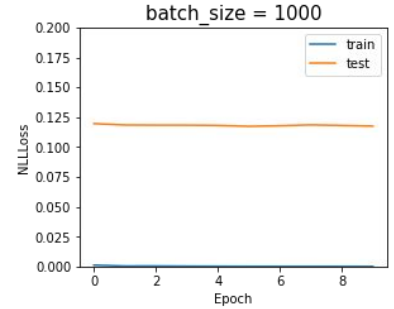
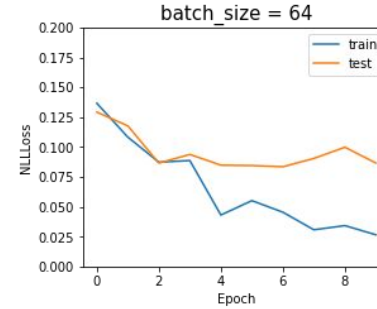
MNIST

MNIST - Canvis a batch size

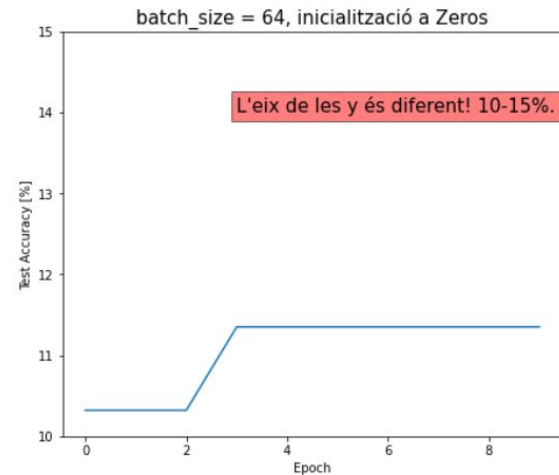
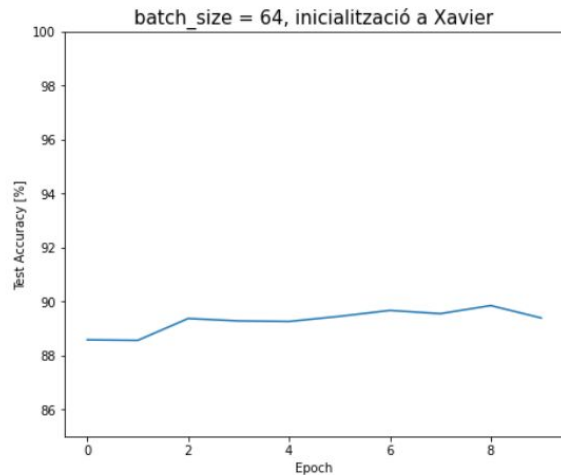
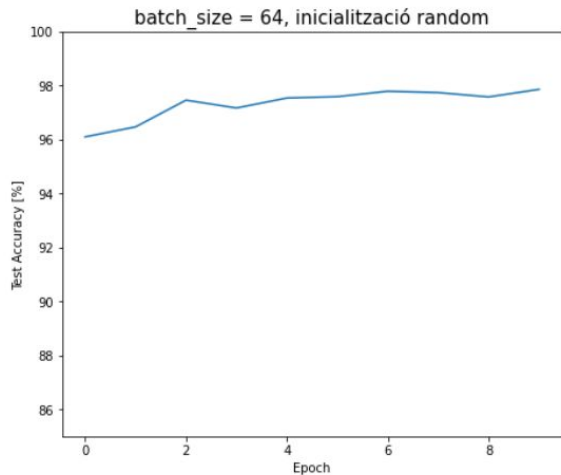
- Ambdós cops amb random initialization.
- Batch Size = 64 i 1000.

```
[18] import datetime
print('Per a batch_size = 64, tenim ' + str(datetime.timedelta(seconds=temps_batchSize64)))
print('Per a batch_size = 1000, tenim ' + str(datetime.timedelta(seconds=temps_batchSize1000)))

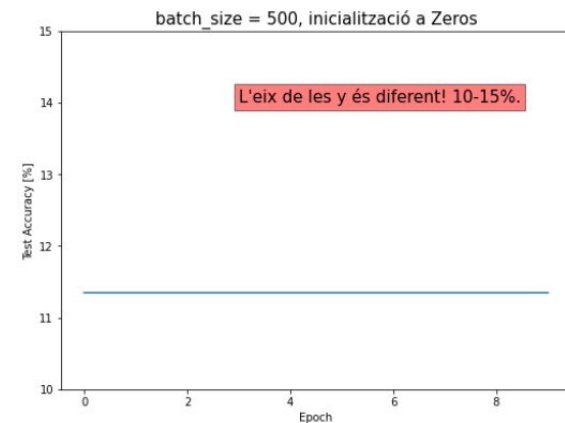
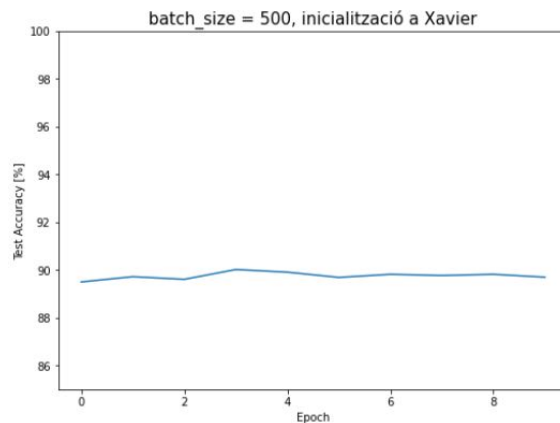
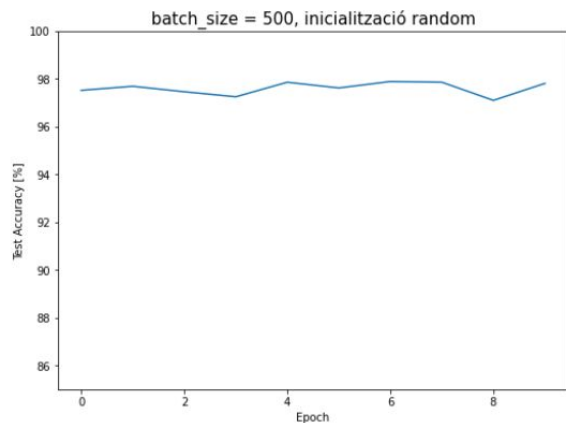
Per a batch_size = 64, tenim 0:02:08.151730
Per a batch_size = 1000, tenim 0:01:53.524381
```



MNIST - Weight Initialization

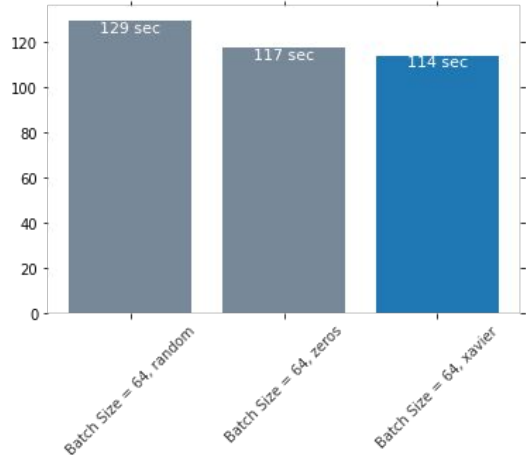


MNIST - Weight Initialization

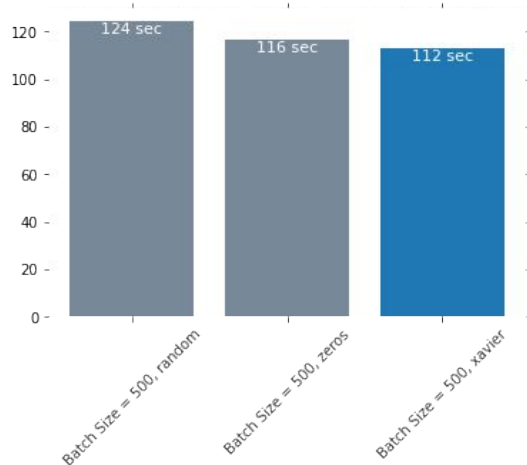


MNIST - Execution times

Execution times when batch size = 64 for MNIST train + test



Execution times when batch size = 500 for MNIST train + test



Per a batch_size = 64, inicialització random, tenim 0:02:09.742058
 Per a batch_size = 64, inicialització a Zeros, tenim 0:01:57.780393
 Per a batch_size = 64, inicialització a Xavier, tenim 0:01:54.111505

Per a batch_size = 1000, inicialització random, tenim 0:01:39.104455

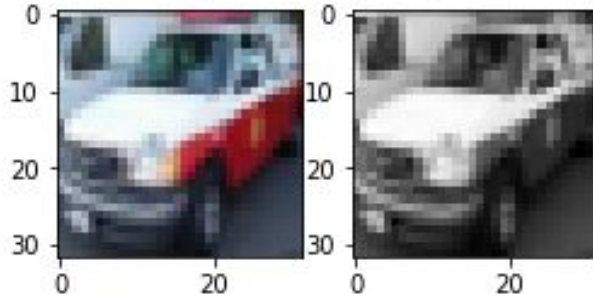
Per a batch_size = 500, inicialització random, tenim 0:02:04.328960
 Per a batch_size = 500, inicialització a Xavier, tenim 0:01:52.805358
 Per a batch_size = 500, inicialització a Zeros, tenim 0:01:56.679245

Per a batch_size = 1, i només una època, tenim 0:01:31.543283
 Multiplicant per a 10 èpoques, equival a 0:15:15.432827

CIFAR10

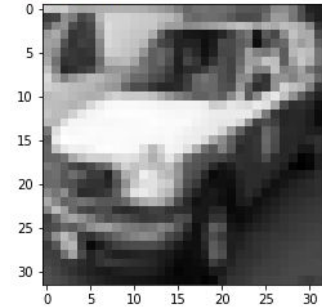
CIFAR10

- NN utilitzada = 1 canal d'entrada.
- CIFAR10 = 3 canals (RGB).
- Error de mides del tensor al entrenar amb 1 canal.



Imatge de 32x32
1 canal (B&W)
Etiqueta = 1 (Car)

Img shape: (32, 32, 1)
Label: [1]



Conclusions

Conclusions

- Batch Size gran (1000):
 - **Overfitting**.
 - Temps d'execució més **ràpid**.
 - Pitjor **precisió** al test.
- Batch Size adequat (64):
 - Millor **precisió** al test.
 - Temps d'execució lleugerament més **lent**.
 - La funció de pèrdues **decreix** conforme fa més iteracions.
- Inicialització de pesos:
 - Millor cas: **random** (precisió ~98%).
 - Pitjor cas: **zeros** (precisió ~11%).

Gràcies per la
vostra atenció