

ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ**1^η Εργασία****Παραδείγματα εφαρμογής**

Τόσο στην εκπαίδευση του MNIST όσο και του CIFAR10 dataset χρησιμοποιώ 100 units στο hidden layer.

A) Κατηγοριοποίηση με το MNIST dataset

Batch-size = {100, 200, 300}

Epochs = { 200, 250}

Lamda = { 0.1, 0.01}

1. Χρησιμοποιώντας την συνάρτηση ενεργοποίησης $h(a) = \log(1+e^a)$ έχουμε:

➤ Για batch-size = 100,

i) epochs = 200, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.8754 = 87.5%

ii) epochs = 200, lamda = 0.01, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.8773 = 87.7%

iii) epochs = 250, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.8803 = 88%

➤ Για batch-size = 200,

i) epochs = 200, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.8804 = 88%

ii) epochs = 200, lamda = 0.01, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.8815 = 88.1%

iii) epochs = 250, lamda = 0.01, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.8860 = 88.6 %

➤ Για batch-size = 300,

i) epochs = 200, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.8869 = 88.7%

2. Χρησιμοποιώντας την συνάρτηση ενεργοποίησης $h(a) = (e^a - e^{-a}) / (e^a + e^{-a})$ έχουμε:

➤ Για batch-size = 100,

i) epochs = 200, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.9234 = 92.3%

ii) epochs = 250, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.9283 = 92.8%

➤ Για batch-size = 200,

i) epochs = 200, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.9246 = 92.4%

ii) epochs = 300, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.9336 = 93.4%

➤ Για batch-size = 300,

i) epochs = 200, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.9362 = 93.6%

ii) epochs = 300, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.9349 = 93.5%

3. Χρησιμοποιώντας την συνάρτηση ενεργοποίησης $h(a) = \cos(a)$ έχουμε:

➤ Για batch-size = 100,

i) epochs = 200, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.9424 = 92.3%

ii) epochs = 300, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.9496 = 95%

➤ Για batch-size = 200,

i) epochs = 200, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.9443 = 94.4%

ii) epochs = 300, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.9525 = 95.2%

B) Κατηγοριοποίηση με το CIFAR-10 dataset.

1. Χρησιμοποιώντας την συνάρτηση ενεργοποίησης $h(a) = \log(1+e^a)$ έχουμε:

➤ Για batch-size = 100,

epochs = 100, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.4246 = 42.5%

➤ Για batch-size = 200,

epochs = 100, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.4236 = 42.4%

➤ Για batch-size = 300,

epochs = 100, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.4200 = 42%?????

2. Χρησιμοποιώντας την συνάρτηση ενεργοποίησης $h(a) = (e^a - e^{-a}) / (e^a + e^{-a})$ έχουμε:

➤ Για batch-size = 100,

epochs = 100, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.4410 = 44.1%?????

➤ Για batch-size = 200,

epochs = 100, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.4236 = 42.4%?????

➤ Για batch-size = 300,

epochs = 100, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.4466 = 44.6%?????

3.Χρησιμοποιώντας την συνάρτηση ενεργοποίησης $h(a) = \cos(a)$ έχουμε:

➤ Για batch-size = 100,

epochs = 100, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.4629 = 46.30%????

➤ Για batch-size = 200,

epochs = 100, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.4662 = 46.62%????

➤ Για batch-size = 300,

epochs = 100, lamda = 0.1, learning-rate = $0.5/N = 8.3e^{-6}$

⇒ Accuracy = 0.4646 = 46.46%????